Appendix VI

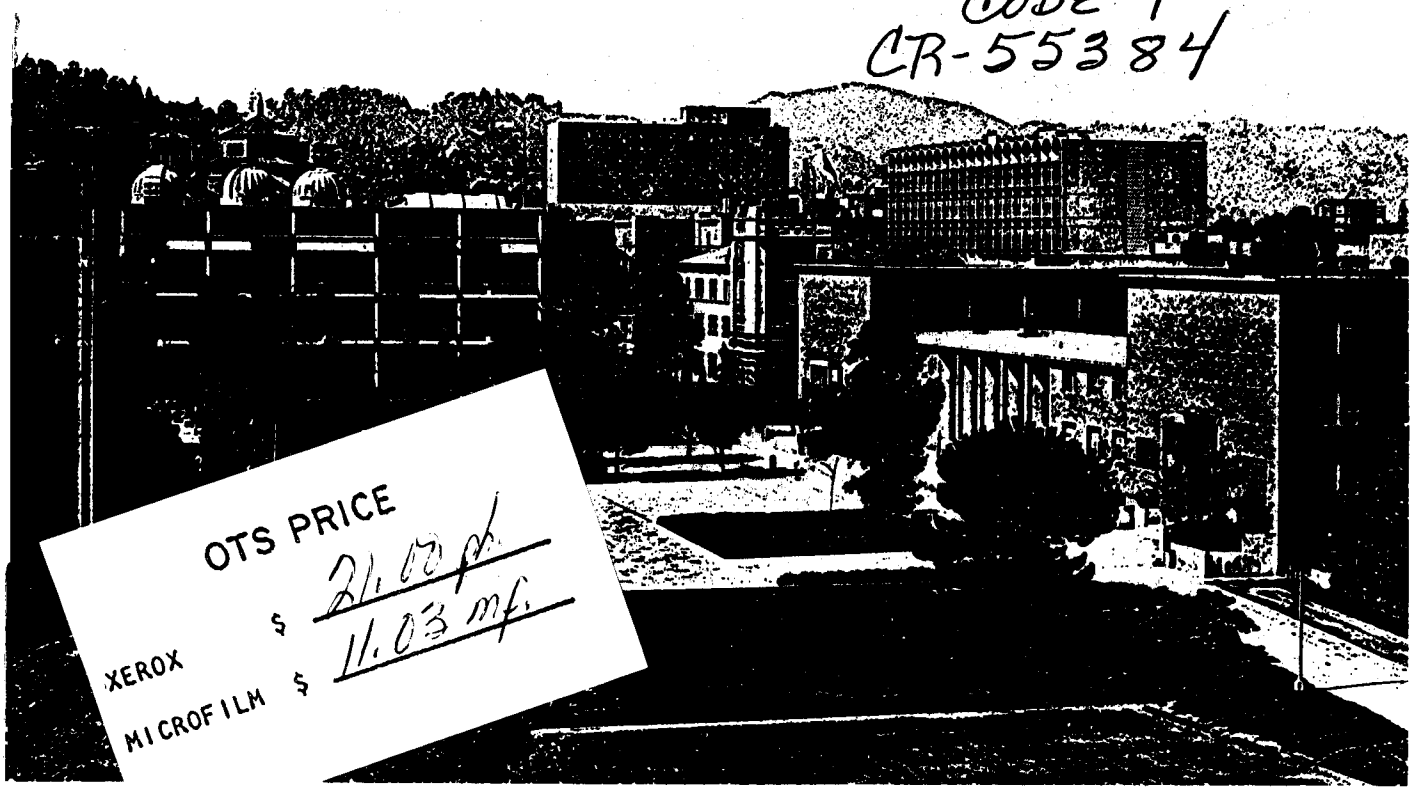
NSG-498

361 p.

r Institute in Planetary Physics

University of California, Los Angeles • 1963

N64-15182–N64-15185
CODE-1
CR-55384



OTS PRICE
XEROX $ 21.00 ph
MICROFILM $ 11.03 mf

GORDON J. F. MACDONALD

ophysics and Planetary Physics

PLANETARY PHYSICS

NOTES FOR THE

SUMMER INSTITUTE IN PLANETARY PHYSICS.

APPENDIX VI

1963

CONTENTS

THE INTERNAL CONSTITUTIONS OF THE INNER PLANETS AND THE MOON

THE ESCAPE OF HELIUM FROM THE EARTH'S ATMOSPHERE

HYDROMAGNETIC WAVES IN SPACE

THE INTERNAL CONSTITUTIONS OF THE INNER PLANETS AND THE MOON

by

Gordon J. F. MacDonald

Inst~~itute~~ of Geophysics
and Planetary Physics
University of California
Los Angeles, ~~California~~

1615004

THE INTERNAL CONSTITUTIONS OF THE INNER PLANETS AND THE MOON

TABLE OF CONTENTS

ABSTRACT 1

I. INTRODUCTION 3

II. INTERNAL CONSTITUTION OF THE EARTH 9

Earth's Gravitational Field 9

Figure of a Fluid Earth 12

Interpretation of Deviations from Equilibrium 18

Interpretation of the Flattening 21

Body Tides and the Rotation of the Earth 22

Body Tides 22

Love Numbers 24

Tidal Variation of Gravity 26

Chandler Wobble 29

Elasticity of the Earth at High Frequencies 31

Propagation of Elastic Waves 32

Determination of the Variation of Density 36

The Earth's Free Oscillations 40

Observations of the Earth's Free Oscillations 44

Structure of the Spectrum 45

Chemical Composition of the Earth 46

Thermal Constitution of the Earth 50

Radioactive Composition of the Earth 51

Surface Heat Flow Measurements 56

Heat Flow and the Deep Distribution of Radioactivity 59

Thermal Conductivity of Silicate Material 60

Melting Relations at High Pressures 64

Convection 68

Calculation of Internal Thermal Conditions 75

Expansion of the Earth 81

Interpretation of the Variation of Seismic Velocities in the Mantle 85

Interpretation of the Low-Velocity Layer 89

Structure of the Mantle between 200 and 900 Kilometers 90

Chemical Composition of the Earth's Core 92

The Earth's Magnetic Field 94

The Origin of the Magnetic Field 98

III. INTERNAL CONSTITUTION OF THE MOON 101

The Moon's Gravitational Field 101

Tides on the Moon 106

Seismic Activity on the Moon 108

Thermal Constitution of the Moon 110

Thermal Radiation from the Moon 110

Calculations on the Thermal History of the Moon 111

The Moon's Magnetic Field 116

IV. INTERNAL STRUCTURE OF MARS 117

Introduction 117

Geometric Figure of Mars 117

Gravitational Field 118

Density Distribution within Mars 121

Thermal Constitution of Mars 122

Magnetic Field of Mars 124

V. THE INTERNAL STRUCTURE OF VENUS AND MERCURY AND THE ROTATION OF THE PLANETS 125

Introduction 125

Rotation of Venus and Mercury 126

Thermal Structure of Venus 127

Interior of Mercury 130

The Magnetic Field of Venus 131

REFERENCES 132

LIST OF TABLES

TABLE I - Estimates of Zonal Harmonics 13

TABLE II - Estimates of Tesseral Harmonics 14

TABLE III - Comparison of Observed Gravity and Gravity Calculated for a Compensated Crust 17

TABLE IV - Maximum Stress Differences Resulting from Deviations of Earth from Equilibrium Fluid 20

TABLE V - Amplification Factor $(1 - \frac{3}{2} k + h)$ and Phase Lag for Body Tides 27

TABLE VI - Distribution of Mass in Crust 35

TABLE VII - Crustal Abundances 48

TABLE VIII - Possible Samples of the Upper Mantle 49

TABLE IX - Half-life and Heat Generation of Important Heat Producing Isotopes 53

TABLE X - Heat Production in Rocks and Chondritic Meteorites 54

TABLE XI - Mean Heat Flow Based on 757 Observations 57

TABLE XII - Radiative Conductivity Calculated from Absorption Spectra 65

TABLE XIII - Coefficient for Simon Equation 67

TABLE XIV - Multipole Components of the Earth's Magnetic Field 96

TABLE XV - Computed Periods of the Spheroidal Oscillations of the Moon 109

TABLE XVI - Radius and Mean Density of Mars 119

TABLE XVII - Thermal Parameters for Mars Models I and III 123

TABLE XVIII - Thermal Parameters for Models of Venus 129

LIST OF FIGURES

	Figure Heading	After Page
Figure 1	Geoid heights in meters referred to an ellipsoid with a flattening of 1/298.3 (after Kaula, 1963b).	18
Figure 2	The variation of compressional (C_p) and shear (C_s) wave velocity and density within the earth. (Gutenberg, 1959).	34
Figure 3	Total elastic energy in the toroidal oscillations. (MacDonald and Ness, 1961).	43
Figure 4	Comparison of the resonant frequencies for the toroidal oscillations in the Gutenberg and homogeneous models of the earth.	43
Figure 5	Elastic energy per unit radius in the toroidal oscillations for the Gutenberg model earth.	43
Figure 6	Power spectrum of the gravity record after the Chilean earthquake of 22 May 1960.	44
Figure 7	Power spectrum of a quiet period one month after the Chilean earthquake.	44
Figure 8	The comparison of the calculated and observed frequencies of the spheroidal oscillations.	45
Figure 9	The heat production per column of unit surface cross-section.	56
Figure 10	Orthogonal function representation of 757 heat flow values (Lee and MacDonald, 1963).	58

	Figure Heading	After Page
Figure 11	Vertical distribution of uranium in g g^{-1} for the models illustrated in Figures 14 - 23.	79
Figure 12	The concentration of potassium in g g^{-1} for the models illustrated in Figures 14 - 23.	79
Figure 13	The initial vertical distribution of temperature used in the models illustrated in Figures 14 - 23.	79
Figure 14	The development of the temperature field in a continent formed 3 x 10^9 years ago.	79
Figure 15	Further development of the temperature field illustrated in Figure 14.	79
Figure 16	The development of the surface heat flow with time for the models illustrated in Figures 14 and 15.	80
Figure 17	The development in time of the thermal field for a continent formed 3 x 10^9 years ago, where the total radioactivity is that of a chondritic earth and the initial temperature is along Curve B of Figure 13.	80
Figure 18	Further development of thermal field illustrated in Figure 17.	80
Figure 19	The development in time of the surface heat flow for the models illustrated in Figures 17 and 18.	80
Figure 20	The temperature field after 0.95 x 10^9 years for a continent formed 3 x 10^9 years ago with initial temperature distribution along Curve A of Figure 13 and with a chondritic mantle radioactivity.	81

	Figure Heading	After Page
Figure 21	Temperature after 1.89 x 10^9 years for a continent formed 3 x 10^9 years ago.	81
Figure 22	The present temperature distribution for a continent formed 3 x 10^9 years ago with an initial temperature along Curve A of Figure 13.	81
Figure 23	The development of surface heat flow for the temperature models shown in Figures 20,21,22.	81
Figure 24	The rate of radial expansion in kilometers per 10^9 years for an earth having an initial temperature distribution along Curve B of Figure 13.	84
Figure 25	The radial expansion for two model earths having initial temperature along Curve B of Figure 13.	85
Figure 26	Stability relations between sillimanite and kyanite (Al_2SiO_5); quartz and coesite; and albite ($NaAlSi_3O_8$), jadeite ($NaAlSi_2O_6$), and quartz.	88
Figure 27	The calculated curve for the breakdown of olivine to periclase plus stishovite.	91
Figure 28	The variation of temperature in a moon having the chondritic radioactivity uniformly distributed.	112

	Figure Heading	After Page
Figure 29	Maximum temperature within the moon as a function of the depth of burial of heat sources.	113
Figure 30	Maximum temperature within the moon as a function of the total radioactivity of the moon.	113
Figure 31	Dependence of surface heat flow on the depth of burial of heat sources.	114
Figure 32	The variation of surface heat flow with total radioactivity.	114
Figure 33	Dependence of surface density and radius of the inner core of Mars as a function of the hydrostatic flattening.	121
Figure 34	The variation of density with depth in Mars for two models.	122
Figure 35	Density models for Mars with outer radius of 3345 km and a flattening of 4.7×10^{-3}.	122
Figure 36	Calculated temperature distribution of Mars assuming uniform chondritic radioactivity.	122
Figure 37	Calculated temperature distribution of Mars for a model with initial temperature twice as high as for the model illustrated in Figure 36 where the chondritic radioactivity is uniformly distributed.	122

	Figure Heading	After Page
Figure 38	The density of angular momentum associated with the rotation for the planets.	125
Figure 39	Calculated temperature distribution in Venus for an initial temperature of 1000°C.	128
Figure 40	Calculated temperature distribution in Venus for an initial temperature of 500°C.	129

THE INTERNAL CONSTITUTIONS OF THE INNER PLANETS AND THE MOON

Gordon J. F. MacDonald
Institute of Geophysics
and Planetary Physics
University of California
Los Angeles, California

ABSTRACT

The internal structures of the moon, Mars, Venus, and Mercury are examined in the light of what is known about the constitution of the earth. The gravitational figure of the earth as obtained from orbits of artificial satellites is used to estimate the possible deviations from hydrostatic equilibrium on other planets. Observations of the orbital and rotational motion of the moon are consistent with the hypothesis that the interior of the moon supports density inhomogeneities of the same order as those supported by the earth. The available data on the moon are insufficient to determine whether or not the moon is differentiated. The orbits of Phobos and Deimos yield an adequate value for the moment of inertia of Mars. The moment of inertia and the mass are consistent with a metallic core containing about 10 per cent of the mass of Mars. The observation of the possible magnetic field of Mars would be of importance both to the understanding of planetary magnetic fields and elucidating the internal structure of that planet. Seismic investigations on the earth yield an equation of state for silicates to pressures of about 1×10^6 bars. This equation of state is used in determining density variation within Mars.

The surface heat flow for the earth is consistent with the hypothesis that the concentration of radioactive elements is the same as that in chondritic meteorites. The observed ratio of potassium to

uranium in surface and near-surface rocks is not consonant with the chondritic hypothesis. The moon can be of chondritic composition only if it is differentiated with the radioactivity concentrated in the upper few hundred kilometers. A chondritic composition for Mars would require a differentiation in excess of that consistent with its mass and moment of inertia. It is concluded that a chondritic composition is not a satisfactory chemical model for the inner planets.

THE INTERNAL CONSTITUTIONS OF THE INNER PLANETS AND THE MOON

I. INTRODUCTION

At the present and in the foreseeable future, the interiors of planets can only be explored through indirect observations. Before the advent of planetary probes, planets other than the earth could be investigated by visual and radio observations. Preliminary exploration of space has already yielded data of considerable importance in the interpretation of the internal constitution of other bodies. As an example, it is now known that neither the moon nor Venus possesses a magnetic field comparable in strength to that of the earth. These observations provide further limits to the speculations regarding both the origin of planetary magnetic fields and the nature of planetary interiors.

Our knowledge of the earth's interior is considerable compared with the information regarding other planets. The data on the internal structure of the earth makes possible a construction of models of the other planets. These models provide quantitative estimates of parameters that can be measured by instruments carried in vehicles that either orbit or land on the moon, Mars, or Venus. Observations obtained in this way can then be compared with the models and the observed deviations will provide both a guide to the internal constitution of the other planets and a test of the understanding of our own planet.

Knowledge of the earth and its interior has grown in the past few years. Jeffreys (1937) first examined the constitution of the earth-like planets in the light of geophysical knowledge. During the past

ten years, the understanding of the internal constitution of the earth has increased. Large-scale variations of the earth's gravitational field are well known as a result of studies of the orbits of artificial satellites. The low-frequency oscillations of the earth provide new data regarding the elastic properties of the earth. Although major problems still abound in the theory of the earth's interior, its internal distribution of mass and elasticity are relatively well determined. The question of the thermal regime remains a pressing problem. The origins of continents and oceans and other surface features are closely tied to the anelastic behavior of the earth and little progress has been made in understanding the behavior of silicate materials when subjected to small stresses for long periods of time. The internal part of the earth's magnetic field undoubtedly originates in magneto-hydrodynamic motions of the fluid core. The driving mechanisms for these motions are uncertain. It is possible that the motions result from thermal inequalities arising from radioactive heat sources or heat released through a precipitating fluid. Alternatively, the core may be stirred by the relative motion of the mantle and fluid core; in this case the field is driven by the precessional torques of the sun and the moon.

Since the internal constitution of the earth provides a guide to the nature of the interior of the other terrestrial planets, we will first review those areas of the study of the earth that have contributed most heavily to the understanding of its internal constitution. The methods, concepts, and techniques applicable to the earth should

also apply to the study of the moon and the inner planets. After considering the state of knowledge about the earth, we will discuss the current state of knowledge of the internal constitution of the moon, Mars, Venus, and Mercury. Particular attention will be paid to those problems about which a great deal will be discovered in the early stages of planetary exploration.

The determination of the earth's figure is a principal concern of geodesy. Since the external figure of the earth is the result of its rotation and internal mass distribution, a detailed knowledge of the earth's external gravitational field sets limits on the internal mass distribution. The departure of the earth's gravitational field from that of a homogeneous sphere results in observable perturbations in the motions of the moon and of artificial satellites. The study of the orbits of artificial satellites has been particularly valuable in providing additional information regarding the earth's external potential.

The gravitational forces of the moon and sun raise tides in the solid body of the earth; and these external tidal forces are precisely known. Accurate observations of the earth's response to these forces yield information on both elasticity and density within the earth and on the deviations from perfect elasticity at the tidal frequencies. The variation of the earth's rotation both in terms of the variation of position of the instantaneous axis of rotation and in terms of the length of day yields information regarding the internal elasticity and anelasticity.

Seismology provides the most powerful tool for the investigation of the earth's interior; the propagation of elastic waves depends on the detailed variation of density and elasticity within the earth. The sources of such elastic waves are for the most part earthquakes though in recent years artificial sources have become increasingly important. The determination of the travel time of these waves permits an estimation of their variation of velocity with the path of travel. This variation yields the radial distribution of velocity. The radial distribution of the velocity of both compressional and shear waves can be combined with astronomical data to obtain the distribution of density within the earth. In recent years an additional powerful investigational tool has been the observation of the free oscillations. The free oscillations involve the planet as a whole and an analysis of the frequencies of the free oscillations reduces the difficulties associated with analyzing many events on different seismographs for several earthquakes.

The interpretation of the variation of elasticity and density within the earth requires a detailed knowledge of the behavior of silicate materials at high temperatures and pressures. In the last few years, great progress has been made in simulating deep-earth conditions in laboratories. The materials that are stable at the earth's surface undergo a vast variety of changes when subjected to conditions that prevail within the upper 100 km of the earth. It is likely that further changes in the structure of matter will occur in laboratory experiments as they more closely approach conditions existing in the

deep interior where pressures are on the order of 10^6 bars and temperatures are on the order of a few thousand degrees. Laboratory data can be combined with seismically derived velocity-depth variation data to construct chemical-physical models of the earth's interior. The laboratory data do not yet cover sufficient range and most models depend heavily on untested theories of the behavior of matter at high temperatures and pressures.

The variation of temperature within the earth remains a matter of great uncertainty. Through the outer few tens of kilometers of the earth, heat is transported by ordinary lattice conduction. Measurements of the surface temperature gradient and conductivity establish that on the average about 64 erg cm^{-2} sec^{-1} flow out of the earth. The surface heat flow is determined by the internal distribution of radioactive heat sources and by the mechanism by which heat is transported. In addition to ordinary thermal conductivity, radiation is effective at high temperatures and there is a possibility that convective motion in solids is the mechanism for the transport of heat. Models of the earth involving the distribution of radioactivity and initial thermal conditions can be constructed. The models give some indication of the possible internal temperature distribution, but the actual distribution remains most uncertain.

After reviewing the state of knowledge of the earth in the fields considered above, we will examine the present state regarding the moon and the other planets. Both for the moon and Mars, astronomical

observations provide constraints on possible internal structures. The situation with respect to Venus and Mercury is much more unsatisfactory since only their mean densities are roughly known.

Observations of the rotation and mean motion of the moon provide data regarding the distribution of density within the moon. These data suggest that the moon is a relatively strong body capable of sustaining substantial stress differences for long periods of time. If the moon is indeed strong, then the internal temperatures must be well below the melting point. These considerations place limits on the possible thermal field within the moon.

The orbital constants of Phobos and Deimos yield the mass and dynamical flattening of Mars. These two quantities can be combined to limit possible models for the internal density distribution. The major uncertainty is the diameter of the planet. Furthermore, the optical figure differs from the dynamical flattening and this difference remains a major problem.

The rotations of the inner planets differ in a suggestive fashion. For the major planets and Mars the density of angular momentum is a simple function of the total mass of the planet. The earth has a deficiency of angular momentum which can be explained in terms of the tidal drag due principally to the moon. Mercury and Venus show a great deficiency in angular momentum density due to tidal interaction with the sun.

II. INTERNAL CONSTITUTION OF THE EARTH

In this section, we review the present state of knowledge regarding the interior of the earth. Those advances made within the last 10 years will be emphasized, as will those methods which are likely to be fruitful when applied to other planets. The data for the earth have been obtained through the use of orbiting artificial satellites, seismometers, gravimeters, magnetometers, and heat flow probes. Similar instrumentation will yield data of first importance on the other planets.

Earth's Gravitational Field

Dynamical geodesy deals with size and shape of the earth and the description of the earth's gravitational field. The results of dynamical geodesy are of great importance to the discussion of the earth's interior. Knowledge of the earth's gravitational field provides limits to the density distribution within the earth and, when combined with the theory of a rotating fluid, yields information regarding the long-term mechanical behavior of the earth.

Prior to the launching of satellites, information regarding the external gravitational potential and the dimensions of the earth was obtained from measurements of arc lengths, determinations of surface gravity, observations of the motions of the moon, and the determination of the lunar-solar precession of the earth (Bomford, 1962). While the astronomical observations are carried out with considerable precision, the earth-bound observations yield data on a planetary scale that are of limited accuracy. The principal reasons for this are the unequal

distribution of observations over the surface and the great variety of instrumental techniques employed in the observations.

Satellites are by no means an ideal geodetic tool, yet they do provide valuable information regarding the large-scale features of the earth's gravitational field. Deficiencies in tracking, incomplete data analysis techniques, and lack of an adequate theory to remove the secular and periodic effects of atmospheric drag all reduce the accuracy with which satellites may be used to determine the external potential of the earth. Since 1958, the detailed reduction of observations of a large number of satellites has led to the determination of 8 zonal harmonics of the earth's gravitational field together with estimates of certain of the tesseral harmonics.

We write the external gravitational potential of the earth in the notation recommended by the International Astronomical Union (Hagihara, 1962).

$$U = \frac{\mu}{r}\left[\, 1 + \sum_{n=1}^{\infty} \sum_{m=o}^{n} \left(\frac{R}{r}\right)^n P_n^{\,m}(\sin \delta) \right. \qquad (1)$$

$$\left. \left(C_n^{\,m} \cos m\lambda + S_n^{\,m} \sin m\lambda \right) \right]$$

μ = $GM_\oplus$ ($cm^3\ sec^{-2}$),

r = distance from center of earth (cm),

R = mean equatorial radius of the earth (cm),

$P_n^{\,m}$ = associated Legendre polynomial,

β = latitude,

λ = longitude.

The coefficients C_n^m and S_n^m are dimensionless quantities specifying the contribution of each harmonic in the external potential. An alternative and useful notation is

$$J_n = - C_n^o \quad . \tag{2}$$

A consistent set of numerical values entering into Equation (1) is

$$\begin{aligned} GM_\oplus &= 3.986\,032 \pm 0.000\,030 \times 10^{20} \text{ cm}^3 \text{ sec}^{-2}, \\ J_2 &= 1082.30 \times 10^{-6}, \\ R &= 6.378\,165 \times 10^{8} \text{ cm}, \\ f &= 1/298.30, \\ \omega &= 7.292\,115\,85 \times 10^{-5} \text{ sec}^{-1}, \\ \gamma_e &= 978.0300, \end{aligned} \tag{3}$$

(Kaula, 1963c; Clarke, 1962; Kozai, 1962a). f is the flattening of the reference ellipsoid, ω is the rate of the earth's rotation with respect to inertial space, and γ_e is the equatorial gravity.

The numerical values listed above combine to yield

$$m = \frac{\omega^2 R}{\gamma_e} \quad ; \qquad m = 3.4678 \times 10^{-3} \quad , \tag{4}$$

for the ratio of the centrifugal to the equatorial acceleration.

The zonal harmonics J_n lead to long-period and secular effects on close satellite orbits. A large number of papers (over 70) have been published listing improved numerical estimates of the J_n. Table I lists a number of the more comprehensive investigations and reviews of the zonal harmonics. Of these, those due to Kozai (1962b) perhaps should be weighted most heavily since they have been derived from an extensive series of observations on several satellites.

O'Keefe and Batchlor (1957) first showed that the sectorial harmonics C_2^2, S_2^2 would give rise to an observable, semi-diurnal oscillation in the secular orbit. The determination of these high-frequency terms in the orbit is subject to far greater uncertainty than the long-period and secular terms. For the most part, Baker-Nunn camera observations have been used. Table II lists estimates of the tesseral harmonics. The most extensive investigation has been carried out by Kaula (1963a,b), who analyzed Baker-Nunn observations for several satellites for periods ranging over several hundred days.

Figure of a Fluid Earth

A rotating fluid develops an equatorial bulge represented in the gravitational potential by finite values of J_n, n even. The difference between the observed J_n for the earth and the J_n for an equivalent rotating fluid is a measure of the deviation of the earth from hydrostatic equilibrium. For a hydrostatic earth (all surfaces of equal

TABLE I

Estimates of Zonal Harmonics

(dimensionless, unit = 10^{-6})

Author	J_2^0	J_3^0	J_4^0	J_5^0	J_6^0	J_7^0	J_8^0	J_9^0
O'Keefe, Eckels, and Squires (1959)	1082.49±0.06	-2.39±0.26	-1.70±0.06	-0.30±0.53	-	-	-	-
Kozai (1961a)	1082.19±0.02	-2.29±0.02	-2.13±0.04	-0.23±0.02	-	-	-	-
King-Hele (1961)	1082.79±0.15	-	-1.4±0.2	-	0.9±0.8	-	-	-
Kaula (1962)	1082.3±0.2	-2.3±0.1	-1.8±0.2	-0.3±0.2	-	-	-	-
King-Hele (1962)	1082.7±0.3	-2.4±0.1	-1.7±0.3	-0.1±0.2	0.8±0.1	-0.4±0.2	-	-
Kozai (1962b)	1082.48±0.06	-2.56±0.012	-1.84±0.08	-0.064±0.019	0.39±0.12	-0.47±0.02	-0.02±0.02	0.12±0.02

TABLE II

Estimates of Tesseral Harmonics

(dimensionless, unit = 10^{-6})

	C_2^2	S_2^2	C_3^1	S_3^1	C_4^1	S_4^1
Kaula (1961)	0.38	-1.64			1.12	-0.26
Kozai (1961b)	0.60	-2.24			-0.25	0.08
Newton (1962)	2.1	-0.8			-1.2	-0.2
Kaula (1963a)	1.19	-1.10	1.91	-0.12		
Kaula (1963b)	1.19	-1.13	1.64	-0.05	-0.31	0.43

density are level surfaces), odd harmonics vanish so that the magnitude of the J_n, n odd, is also a measure of the deviation of the earth from fluid equilibrium.

The coefficient J_2 can be written in terms of the principal moments of inertia C and A as

$$J_2 = \frac{C - A}{MR^2} \quad . \tag{5}$$

Because of the flattening of the earth, the attraction of the sun and moon on the earth's bulge induces a motion of the earth's rotational axis about the pole to the ecliptic. From observations of the precession of the equinoxes, it is possible to obtain a value of $H = \frac{C - A}{C}$ of 1/305.3 where the principal uncertainty is in the mass of the moon (Jeffreys, 1959). The precessional constant H can be combined with the dipole moment J_2 to yield a good estimate of the polar moment of inertia C

$$\frac{J_2}{H} = \frac{C}{MR^2} \quad . \tag{6}$$

The theory of rotating fluids yields a result between hydrostatic flattening f_H and the principal moment of inertia C

$$\frac{C}{MR^2} = \frac{2}{3}\left\{ 1 - \frac{2}{5}\left[\frac{5m}{2f_H} - 1 \right]^{1/2} \right\} \quad , \tag{7}$$

(Jeffreys, 1959). The hydrostatic flattening f_H can then be used to obtain a value for the hydrostatic J_2'

$$J_2' = \frac{2}{3} f_H \left(1 - \frac{1}{2} f_H\right) - \frac{1}{3} m \left(1 - \frac{3}{2} m - \frac{2}{7} f_H\right) \qquad (8)$$

The theory carried to higher orders yields relations between the hydrostatic value of the flattening and the hydrostatic coefficients J_n'.

The deviation of the earth from hydrostatic equilibrium is illustrated in Table III where in the first row, the nonhydrostatic coefficients in the expansion of the earth's potential are listed. These are obtained by subtracting the values of the potential of a fluid earth from the observed values listed in Tables I and II.

The nonhydrostatic components of the earth's potential can be referred to more familiar quantities such as the corresponding gravity anomalies or deviations of the surface of equal potential (geoid) from a standard spheroid of revolution. The gravity anomalies δg are given by

$$\delta g = -\frac{2U_o}{r} - \left(\frac{\partial U}{\partial r}\right)_o \quad , \qquad (9)$$

where the subscript o implies that the potential and its derivatives are evaluated at the surface of the geoid. If the gravity anomaly is expanded in a series of spherical harmonics, the coefficients, δg_n^m, are related to the coefficients in the expansion of the potential by

$$\delta g_n^m \simeq (n-1)\, \gamma_e \begin{pmatrix} C_n^m \\ S_n^m \end{pmatrix} \quad . \qquad (10)$$

TABLE III

Comparison of Observed Gravity and Gravity Calculated for a Compensated Crust

Harmonic	J_2	$-\left[C_2^{\,2} + S_2^{\,2}\right]^{1/2}$	J_3	$-\left[C_3^{\,1} + S_3^{\,1}\right]^{1/2}$	J_4	J_5	J_6	J_7	J_8	J_9
Non-hydrostatic (dimensionless, unit = 10^{-6})	12.0	-1.6	-2.6	-1.9	0.5	-0.06	0.39	-0.47	-0.02	0.12
Deviation of geoid ($\delta z_n^{\,m}$) from equilibrium figure (meters)	-86.4	10.2	16.5	12.1	-3.2	0.38	-2.5	3.0	0.13	-0.76
Gravity anomoly ($\delta g_n^{\,m}$)(milligals)	-11.8	1.6	5.1	3.7	-1.5	0.23	-1.9	2.8	0.18	-0.94
Gravity anomoly expected from an isostatic standard crust (milligals)	0.19	-0.16	-0.33	-0.63	0.35	-2.0	0.98	-2.2	-0.48	

Similarly, the deviation of the geoid from the reference spheroid can be expanded in spherical harmonics and the coefficients, δz_n^m, are related to coefficients in the expansion of the potential by

$$\delta z_n^m \simeq R \begin{pmatrix} C_n^m \\ S_n^m \end{pmatrix} . \qquad (11)$$

Table III lists the deviation of the geoid from the equilibrium figure where the units are in meters. The corresponding gravity anomalies are also given. It is seen that the largest deviation is in the flattening of the earth and amounts to 76 meters in the geoid or 12 milligals in the gravity field (1 gal = cm sec^{-2}). The geoid derived by Kaula (1963b) is shown in Figure 1.

It might be supposed that the observed nonhydrostatic contribution to the external potential could arise from the inequalities in mass represented by the continent-ocean system. The gravity anomalies expected for an isostatically compensated standard crust are listed in Table III. For the low-order harmonics, the gravity anomaly expected from the crust is small compared with the observed gravity anomaly and opposite in sign. This implies that the deviations of the figure of the earth from an equilibrium figure cannot be attributed to the near-surface distribution of mass but must result from inhomogeneities in the density field at depth (Munk and MacDonald, 1960b).

Interpretation of Deviations from Equilibrium

The interpretation of the deviation of the earth's potential from that expected for an equivalent rotating fluid raises the problem of

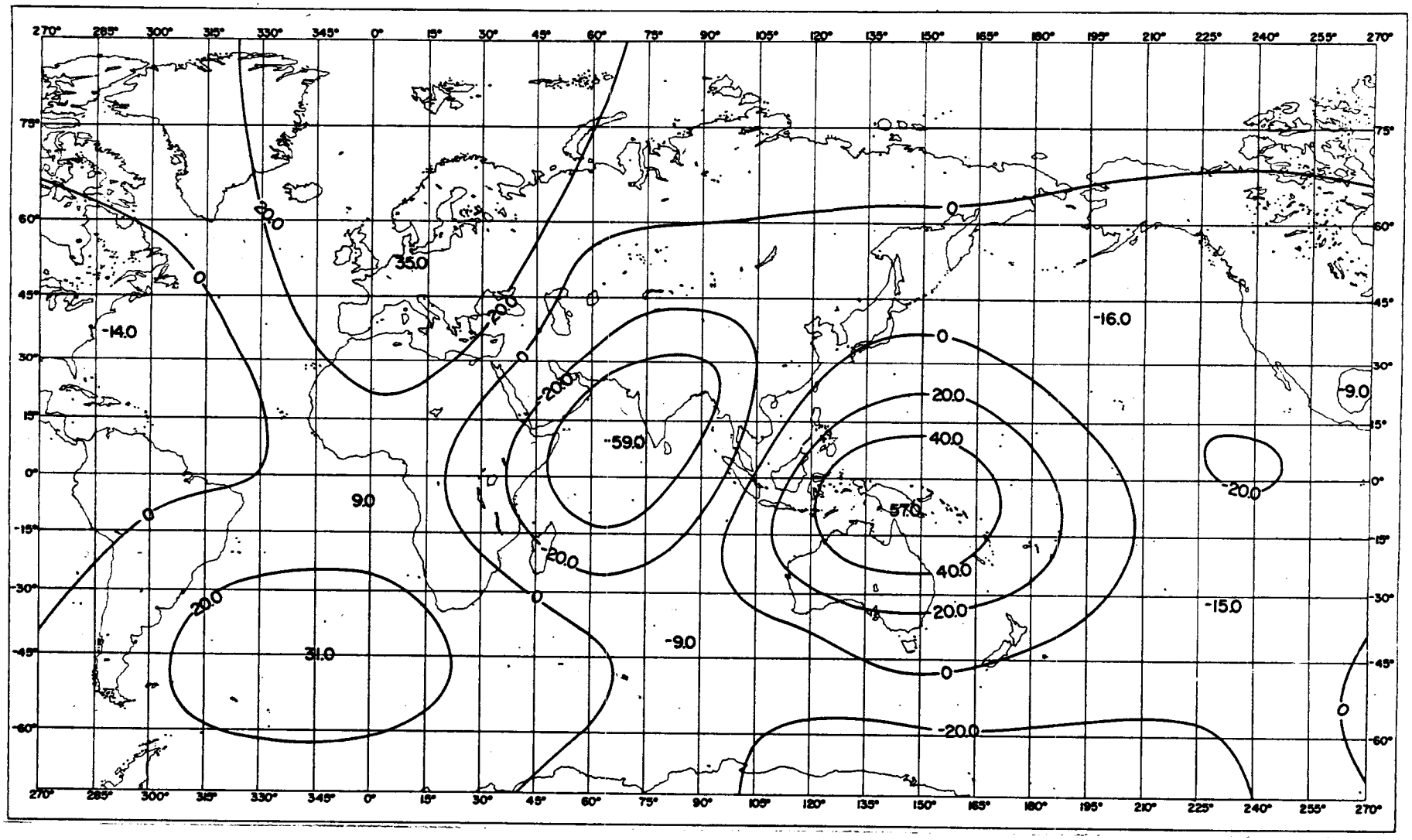


Figure 1: Geoid heights in meters referred to an ellipsoid with a flattening of 1/298.3 (after Kaula, 1963b). The geoid is derived from Baker-Nunn camera observations of satellites 1959 α 1, 1959 η , 1960 ι 2, and 1961 δ 1.

the long-term anelastic response of the deep earth to small stresses. The question of the earth's response to large-scale, small-amplitude stress differences imposed for long periods of time has plagued geophysicists for a century. Kelvin assumed that the earth could be treated as an elastic body even for long-period deformations whereas Darwin proposed that the earth behaves plastically. Today, proponents for both points of view can be found; the gravitational field, as determined from satellite orbits, demonstrates that the earth is, at present, not in equilibrium. Whether it is continuing to adjust to reach equilibrium, or whether the stress differences implied by the inequalities of mass are maintained for long periods of time is an open question.

The simplest interpretation of the density anomalies within the earth supposes that the anomalies are supported by the finite strength of the rocks making up the mantle. The anomalies lead to stress differences, the magnitude of these differences depend on the radial distances over which the stress differences are supported. Kaula (1963d) has evaluated the stress differences implied by the external gravity field for an elastic mantle. Table IV lists the maximum stress difference for the harmonic coefficients given in Table III. An alternative to the finite strength interpretation for the density of anomalies is that the density inhomogeneities are maintained dynamically by convective currents. We will later consider the evidence for convection within the mantle.

TABLE IV

Maximum Stress Differences Resulting from Deviations of Earth from Equilibrium Fluid (Kaula, 1963c)

Harmonic	Maximum Stress Differences in Bars
C_2^0	163
C_2^2	70
S_2^2	65
C_3^0	32
C_3^1	60
S_3^1	5
C_4^0	18

Interpretation of the Flattening

Munk and MacDonald (1960a,b) note that the deviation of the observed bulge from an equilibrium bulge is that expected from the deceleration of the earth. At present, the earth is slowing down at a rate of about 2 seconds per 10^5 years. As the earth slows down, the equilibrium figure changes. If the earth behaves as a viscous fluid, then the earth will adjust toward the equilibrium figure with the time constant set by the viscosity. Alternatively, the earth may behave as a planet with a finite strength. Stresses due to the changing rotation then build up until the finite strength is reached. At this point stresses can be relieved either by a continual adjustment with fracture and flow at the level of finite strength or by a single, catastrophic readjustment to the equilibrium figure.

At the present rate of deceleration, the nonequilibrium bulge of the earth would develop in about 10^7 years. On the viscous fluid interpretation, the time constant combines with a rigidity of 10^{12} dyne cm^2 to yield a value for the viscosity of 10^{26} c.g.s. Alternatively, on the finite strength hypothesis, the mantle has a creep strength of about 200 bars.

The inequalities in figure, other than that associated with the bulge, also lead to substantial stress differences. The origin of these inequalities is uncertain. They may represent initial inhomogeneities in the density distribution. Alternatively, they may be associated with the process of continent formation in which the

vertical redistribution results in density anomalies at depth. This latter interpretation is strengthened by the fact that the low-order harmonics are correlated with the gravity field in a way opposite to that expected from an isostatically compensated crust.

Body Tides and the Rotation of the Earth

Data on the earth's gravitational potential both provides information regarding the distribution of density within the interior and, by comparing the observed values with those of an equivalent rotating fluid, yields data on the earth's anelastic behavior. Because of the large scales involved, it is safe to assume that the gravitational anomalies are associated with anelastic responses of a long time-scale. As we will review later, data from seismology provides information on the anelastic response of the earth to periods on the order of hours and seconds. The earth also yields to other distorting forces having periods intermediate between the very long secular changes and the high-frequency seismic deformations.

Body Tides

The attraction of the moon and the sun produces deformations in the body of the earth analogous to the tides in the oceans. The attraction of an external body on a spherical earth may be considered to be of two parts. First, there is a net force attracting the earth toward the external body in question which is of such a magnitude it would seem that all the earth's mass were concentrated at the center of gravity. Second, there is a differential force on the various

parts of the earth's volume because they are at slightly different angular positions from the center of the sun or moon. These differential or tide-raising forces produce the ocean tides. They also deform the solid earth, tending to elongate it into an ellipse whose major axis is directed toward the disturbing body. The ocean tides are large and more obvious, but the deformation of the solid earth also produces observable effects. These effects yield information on the interior of the earth. The earth tides yield readily to mathematical treatment, whereas ocean tides do not. The free periods of the elastic vibrations in the earth are on the order of an hour and an equilibrium theory is, therefore, applicable to the slowly varying tidal deformations. In the ocean, the free periods are probably on the order of a few days so that they cannot be treated as equilibrium phenomena. As we shall see, the ocean tides complicate the analysis of body tides.

In addition to the forced motions induced by the action of the sun and moon, there is a free nutation of the earth's axis of rotation. Euler showed that a free body symmetrical about its axis of rotation would rotate permanently about the axis of greatest moment. A small disturbance makes the axis move in a cone about the axis of greatest moment and it completes its revolution in a period of $\frac{C}{C - A}$ times the period of rotation. The wobble in the polar rotation should show up in astronomical observations as a change in latitude with a 305-day period. During the 19th century, a number of astronomers searched

for changes in latitude with such a period. The results were inconclusive. Chandler (1891) announced a term in the latitude variation with a period of 428 days, 40 per cent longer than Euler's classic value. Later, Newcome was able to demonstrate that the yielding of the solid earth and the oceans should bring an increase in the period from 10 to 14 months. He attributed one-fourth of the increase to the mobility of the oceans, the remainder to the earth's elasticity. The lengthening of the period can be understood as follows: The rotation produces an elastic strain which is symmetrical about the instantaneous axis and does nothing to displace it. This is superimposed on a permanent flattening, the two together making up the total flattening. The permanent part affects the axis of rotation just as such a condition would affect a rigid body and it is this part that determines the period; however, both parts are attracted by the sun and moon and contribute to the precession. The lengthening of the period from 10 to 14 months thus gives a measure of the elastic yielding of the earth. In addition, the Chandler wobble is not of constant amplitude but decays. If a source is given, the rate of decay provides information regarding the deviations from elasticity within the earth.

Love Numbers

A description of the deformation of the earth due to both the tide-raising forces of the sun and the moon and the 14-month wobble is best given in terms of certain numbers introduced by Love (1927). We consider the earth's response to a disturbing potential U(R) described in terms of a spherical harmonic of degree two; the tidal

forces of the moon and the sun and centrifugal forces arising from the earth's rotation can be written as gradients of such a potential U. The resulting deformation defines the Love numbers h and k as follows: The ground is lifted by an amount h U/g, where g is the surface value for the gravitational acceleration. The additional gravitational potential at the displaced surface arising solely from this redistribution of mass is kU. The numbers, h and k, are Love numbers of degree two. If the disturbing functions require higher order harmonics for their description, then higher order Love numbers would be defined. The factor, 1 + k, is a factor allowing for the attraction of the bulge by itself and the response of hU/g takes this self-attraction into account. A fluid covering the globe at the surface would remain equi-potential and be lifted by (1 + k)U/g relative to the center of the earth and by (1 + k - h)U/g relative to the sea-bottom. In addition to the vertical displacement of the solid surface by hU/g, there is a horizontal displacement with component

$$\frac{\ell}{g}\frac{\partial U_2}{\partial \beta}\quad , \qquad (12)$$

$$\frac{\ell}{g}\frac{1}{\cos\beta}\frac{\partial U_2}{\partial \lambda}\quad ,$$

where the third Love number, ℓ, describes the horizontal yielding of the surface.

The Love numbers are dimensionless parameters which neatly summarize some of the earth's elastic properties. The determination of the second-order Love numbers fixes the mean elasticity of the planet.

Tidal Variation of Gravity

The tidal deformation displaces the surface and its acceleration to first order at the spherical surface. r = R is

$$g\left[\, 1 - \frac{2U}{Rg}\left(1 - \frac{3}{2}k + h\right)\right] \quad . \tag{13}$$

Gravity measured at the surface of the earth will show a tidal variation. The effect of the yielding of the earth is to amplify the gravity by the factor $(1 - \frac{3}{2}k + h)$.

A large number of observations of the variation of gravity have been carried out by observatories around the world. Of particular interest is a set of observations made by Harrison, et al., (1963) using a single instrument at 13 locations. The results are summarized in Table V. The Love number, k, can be derived from the observed $(1 - \frac{3}{2}k + h)$ if some relation is assumed between k and h. Calculations based on an assumed elastic model for the earth show that $h \approx 2.04$ (Molodenskii, 1961) and that this relation does not depend in a sensitive way on the details of the internal elasticity. A value of 1.184 for $(1 - \frac{3}{2}k + h)$ yields k = 0.34. This can be compared with k = 0.25 obtained by Pariisky (1961, 1963) for central Asia.

Table V shows an important variation in the amplification of the tide and the phase lag with position of the observing station. It seems likely that ocean tides are at least partially responsible, particularly for the observations in coastal stations. Ocean tides,

TABLE V

Amplification Factor $(1 - \frac{3}{2} k + h)$ and Phase Lag for Body Tides
(after Harrison, et. al, 1963)

Tide	O_1	P_1K_1	M_2	S_2K_2
Period (hour)	25.83	23.92	12.42	12.00
		CONTINENTAL-STATIONS		
New Delhi	1.132-1.4°	1.116+1.2°	1.153+1.6°	1.153+1.4°
Bunia			1.185+0.8°	1.185+1.6°
Trieste	1.137+3.2°	1.135+1.6°	1.178+4.6°	1.135+4.8°
Lwiro			1.161-1.8°	1.181-0.7°
		OCEANIC STATIONS		
Honolulu	1.181-6.1°	1.130-3.6°	1.190-2.0°	1.200-4.6°
Glendora	1.221-1.2°	1.186-4.3°	1.184+3.8°	1.258+1.3°
Wake	1.315+1.6°	1.231+2.9°	1.174-0.5°	1.199-0.8°
Manila	1.205+6.7°	1.121+4.1°	1.203+2.6°	1.207+1.9°
Saigon	1.088+10.8°	1.030+11.5°	1.163+0.5°	1.154-1.4°
Bidston	1.118-0.5°	1.141-0.4°	1.155-1.7°	1.183-3.6°
Bermuda	1.198-0.3°	1.133-0.9°	1.249+5.2°	1.174+3.4°
Winsford	1.197+1.4°	1.153 0.0°	1.215-4.1°	1.189-1.1°
MEAN	1.179+1.4°	1.138+1.2°	1.184+0.8°	1.185+0.2°

unlike body tides, are not an equilibrium phenomenon and, since they differ in phase from the disturbing potential, the local value of acceleration is perturbed. The variations among the continental stations indicate that blocks of the earth are moving relative to each other. This observation is of interest since it may provide a clue as to the mechanism by which tidal energy is dissipated.

The magnitude of the torque which arises from the lunar tides and acts to retard the earth's rotation may be estimated from the acceleration of the moon in its orbit. On the basis of astronomical observations during the last 200 years, Munk and MacDonald (1960a) derive a value of 3.9×10^{23} dyne cm for this torque. The magnitude of the torque will depend on the phase angle by which the bodily tide lags behind the passage of the disturbing body. The best determinations of phase lags are those for M_2, the mean of which of 0.8 (Table V) corresponding to a time lag of slightly less than two minutes. This lag occurs in the total gravity tide which is the sum of the direct attraction of the moon and the effects of the earth bodily tide. There is no phase lag in the former contribution and the latter contributes only a fraction $(h - \frac{3}{2} k)/(1 - \frac{3}{2} k + h)$ of the total gravity tide. Employing an amplification factor of 1.186, the phase lag is 2.6 degrees. This corresponds to a tidal torque of 4.6×10^{23} dyne cm. This torque is consistent with the observed value of 3.9×10^{23} dyne cm taking into consideration the difficulties in measuring the small phase lag.

The phase lag indicates the magnitude of the rate of energy dissipation within the earth but does not identify the sink of the energy. Dissipation could take place either as a result of the interaction of the ocean tides or of the imperfections from perfect elasticity within the body of the earth. However, the variability of amplification factors and variations of phase lag suggest that some proportion of the energy must be dissipated in the frictional interaction of crustal blocks.

Chandler Wobble

Observations of the latitude of the stations belonging to the International Latitude Service determine the wobble of the instantaneous axis of rotation. A spectral analysis of the variation of latitude shows two principal peaks, one with a period of a year and another with a period of 429 days (Munk and MacDonald, 1960a). The lengthening of the period over that of an Eulerian rigid body is due both to the elastic yield of the earth and to the fluid response of the core and ocean.

The period of wobble can be interpreted in terms of a Love number k; $k = 0.29 \pm 0.01$. In interpreting this value, it is important to realize that the earth as a whole is involved, including both the oceans and the fluid inner core. The oscillations excited by the wobble are only in part communicated to the liquid core so that the full inertia of the core is not involved in determining the period. Indeed, if this core were a perfect fluid occupying a spherical cavity

in a rigid shell, the core would not partake in the wobble at all. The fraction of the core involved in the wobble depends on a flattening, density distribution, and viscosity of the core and on the electromagnetic interaction between the core and mantle. Only the variations in the ellipticity and density due to the finite elasticity of the shell have been investigated. The subtraction of the inertia of the core reduces the period of the Chandler wobble as compared to a static model in which the earth is taken to be an elastic body. The reduction is in the ratio of the inertia of the core to the inertia of the core plus shell. For a rigid outer body, the ratio is 1/10 and the reduction in period is 30 days.

In quite the same way, an inviscid ocean covering the whole earth would shorten the period though only by a very small amount since the ocean's moment of inertia would be minute. If, on the other hand, this ocean remained always aligned with respect to the wobbling axis, then the ocean would lengthen the period. The lengthening due to such an equilibrium tide would be 33 days. If, without any real evidence, we assumed an equilibrium pole tide, then the value of the Love number would change from 0.29 to 0.23. Jeffreys and Vicente (1957) show that the effects of the core increase k by about 0.08. The Love number due to elastic deformations would then be 0.31. This can be compared with 0.34 derived from the tidal variation of gravity. Combined with the gravitational amplification factor of 1.185, this value of k yields a Love number h = 0.64.

The outstanding problems in the interpretation of the Chandler wobble are the excitation of the motion and the damping. Statistical studies of the latitude observations indicate that the damping is high with a Q on the order of 10 - 30. The Q is lower than that obtained from observations of the elastic behavior of the earth at higher frequencies and the reason for the discrepancy is not understood. The damping may be due either to frictional interaction with the ocean or with the fluid core.

The Chandler motion is a free motion and, with the damping, requires excitation. The mechanism of excitation is not known. Munk and Hassan (1961) show that the atmosphere is incapable of giving rise to the observed amplitude of the Chandler wobble. Recent studies by MacDonald (unpublished) indicate that the Chandler wobble may arise through a quadratic interaction between the yearly oscillation and oscillations having a frequency of 0.15 cycles per year. The indicated nonlinearity hints at the complexity of interpreting even such small-amplitude motions as the Chandler wobble.

Elasticity of the Earth at High Frequencies

The knowledge of the earth's interior has been derived almost entirely from the study of the propagation of elastic waves through the earth. This point needs emphasis when the possible exploration of other planets is considered. While other methods yield data regarding the gross characteristics of the interior, only studies of the propagation of seismic waves can provide the detailed variation of the elastic parameters within a planet.

Observation and analysis of the earth's free oscillations provide another method for determining the general features of the variation of seismic velocity with depth. The oscillations involve the earth as a whole; many of the difficulties associated with analyzing several earthquakes on different seismographs are circumvented.

Propagation of Elastic Waves

In an infinite isotropic elastic solid, two types of elastic waves are possible; a compressional or P wave in which the displacement is in the direction of propagation and a shear or S wave in which the displacement is transverse to the direction of propagation. Provided there are no deviations from perfect elasticity, the waves are propagated without change of form and with velocities which depend only on the properties of the medium. The velocities, C_P and C_S, of the P and S waves are

$$C_P = \left(\frac{k + \frac{4}{3}\mu}{\rho} \right)^{1/2} , \qquad (14)$$

$$C_S = \left(\frac{\mu}{\rho} \right)^{1/2} ,$$

where,

k = incompressibility (dyne cm^{-2}),

μ = rigidity (dyne cm^{-2}),

ρ = density (g cm^{-3}).

The ratio of the two velocities is

$$\frac{C_P}{C_S} = \left[\frac{2(1-\sigma)}{1 - 2\sigma} \right]^{1/2} , \qquad (15)$$

where σ is Poisson's ratio. For most solids, σ is about 0.25 and the ratio of the velocities is

$$\frac{C_P}{C_S} \approx 1.7 . \qquad (16)$$

If an elastic wave meets a discontinuity, it is partially transmitted and partially reflected. The interaction of the elastic wave with the boundary changes its character so that a P wave meeting a boundary will be reflected and transmitted partly as P and partly as S. The P and S waves are transmitted anywhere in the solid medium. Near the surface of the body, there are two waves possible. One is the Rayleigh wave in which the motion falls off exponentially in going into the body and which travels at a velocity somewhat less than the velocity of the S wave. The particle motion in a Rayleigh wave is an ellipse in the plane of propagation. Surface waves in which the displacement is at right angles to the plane containing both the direction of propagation and the normal to the surface are called Love waves.

If an elastic solid is sufficiently isolated from its surroundings, reflections from boundaries may set up standing waves. Lamb showed that the vibrations of an elastic sphere can be classified into two groups. In the toroidal or torsional oscillation, a particle executes motion on the spherical surface; there is no radial component of motion. The toroidal oscillations unite to form the horizontally-polarized S

waves at high frequencies and Love waves at lower frequencies. The spheroidal oscillations combine both tangential and radial motions to produce compression and rarefaction. A degenerate spheroidal oscillation involves only radial motion; the entire sphere expands and contracts. The spheroidal oscillations can be combined to describe Rayleigh waves.

Numerous studies of the variation of seismic velocity with depth have been carried out. The distribution obtained by Gutenberg (1959) is shown in Figure 2.

Under continental areas, the velocity in the upper few tens of kilometers varies markedly with depth and location. The velocity of compressional wave ranges from 5 to 7 km sec^{-1}. At a depth which averages some 30 - 35 km, there is a discontinuity in the seismic velocity and below the discontinuity the velocity is usually about 8 km sec^{-1}. The material above this discontinuity, first recognized by Mohorovicic, is termed the crust. In oceanic areas the discontinuity is overlain by 5 - 6 km of volcanic and sedimentary rocks and some 5 km of water. The distribution of mass in a standard crustal section is shown in Table VI (Worzel and Shurbet, 1954).

The region below the crust, termed the mantle, extends to a depth of 2898 km. At this depth, there is a major discontinuity. Below this depth shear waves are not transmitted and at 2898 km the velocity of the compressional waves undergoes a discontinuity. The region

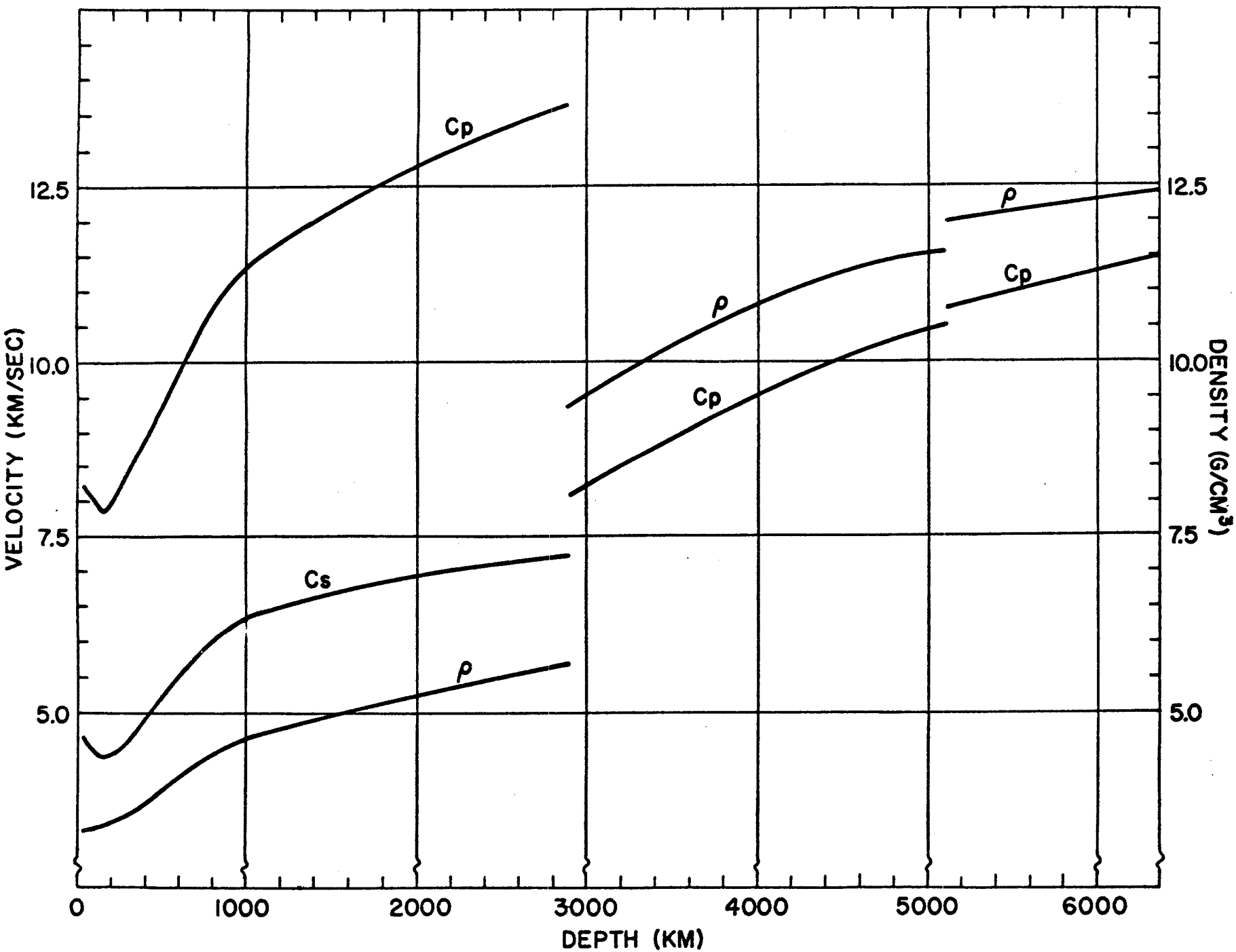


Figure 2: The variation of compressional (C_P) and shear (C_S) wave velocity and density within the earth (Gutenberg, 1959).

TABLE VI

Distribution of Mass in Crust
(after Worzel and Shurbet, 1954)

Oceanic Crust			Continental Crust		
Unit	Density ($g\ cm^{-3}$)	Thickness (km)	Unit	Density ($g\ cm^{-3}$)	Thickness (km)
Water	1.03	5	Crust	2.84	33
Sediment	2.30	1			
Volcanic Rock	2.84	4.5			
Mantle	3.27	22.5			

between 2898 km and the center of the earth is termed the core. The core is divided into an outer fluid region and an inner portion in which the velocity of compressional waves is higher. It is generally assumed that the inner core is solid though no shear waves transversing this part of the earth have been observed.

In the mantle, the velocity, on the average, initially decreases with depth, reaching a minimum at some 150 - 200 km and then rapidly increases to a depth of 900 km. Below this depth, the increase is more gradual. The existence of two velocities throughout the mantle is significant and suggests a normal type of elasticity. The absence of shear waves below the mantle is the principal reason for believing that the outer core is liquid. The absence of shear waves in this region and the drastic decrease of compressional wave velocity are the outstanding seismic features which distinguish the core from the mantle. The surface of separation of core and mantle is usually considered to be sharp and is the best located as well as the largest discontinuity in the earth's interior. At a depth of 5121 km, another discontinuity is encountered. The details of the velocity distribution within the inner core are unknown.

Determination of the Variation of Density

The observations of the earth's gravitational field show that it is very nearly a spheroid of revolution. The mean density of 5.52 g cm^{-3} is about twice that of surface rocks. The greater mean density for the earth as a whole can in part be accounted for by ordinary

compression coupled with phase changes in the silicate materials, but it also implies a concentration of the denser components.

The mean density provides an indication of the extent of the concentration of mass toward the earth's interior. A better measure is the moment of inertia. Observations on artificial earth satellites, coupled with a determination of the precessional constant H, yields a value for the polar moment of inertia (see Section II). If M is the mass and R the equatorial radius, then the polar moment of inertia about the axis of rotation C is equal to $0.3318\ M\ R^2$. For a homogeneous earth, we should have $C = 0.4\ M\ R^2$. The moment of inertia is, therefore, less than that of a homogeneous sphere of the same density and this is strong evidence for a concentration of mass toward the interior.

Suppose that the only stress acting within the interior is hydrostatic pressure p. Suppose further that there is a range of depth in the interior over which the material is chemically homogeneous and where no change of phase takes place. Then the pressure gradient and the density are connected by

$$\frac{dp}{dr} = -\rho g \quad , \tag{17}$$

where g is the gravitational acceleration. The adiabatic bulk modulus K_S is given by

$$K_S = \rho(C_P^2 - \frac{4}{3} C_S^2) \quad , \tag{18}$$

and, therefore, the adiabatic change of pressure with density is

$$\left(\frac{\partial p}{\partial \rho}\right)_S = C_P^2 - \frac{4}{3} C_S^2 \quad . \tag{19}$$

The density will also vary with temperature. The temperature variation may be written

$$\frac{dT}{dr} = \frac{T\alpha}{\rho c_p} \frac{dp}{dr} - \tau \quad , \tag{20}$$

$$= - \frac{T\alpha g}{c_p} - \tau \quad ,$$

where,

α = coefficient of thermal expansion,

c_p = specific heat at constant pressure,

τ = difference between actual gradient of temperature and adiabatic gradient which is $- \frac{T\alpha g}{c_p}$.

Employing the thermodynamic relations between isothermal and adiabatic incompressibility, the equation for density change within the earth may be written in the form

$$\frac{d\rho}{dr} = - \frac{g\rho}{\phi} + \alpha\rho\tau \quad , \tag{21}$$

where

$$\varphi = \frac{K_S}{\rho} = C_P^2 - \frac{4}{3} C_S^2 \quad . \tag{22}$$

The usual procedure in constructing the density curve within the earth is to integrate Equation (21) neglecting the variation of density with temperature.

The variation of gravitational acceleration with radius is determined by

$$g = - \frac{\partial U}{\partial r} \quad , \tag{23}$$

where U is the gravitational potential which satisfies Poisson's equation

$$\nabla^2 U = - 4\pi G\rho \quad . \tag{24}$$

The variation of pressure is

$$\frac{1}{r^2} \frac{d}{dr} \left(\frac{r^2}{\rho} \frac{dp}{dr} \right) + 4\pi G\rho = 0 \quad . \tag{25}$$

Provided the earth were adiabatic and homogeneous in phase and chemical composition, Equations (21) and (25) together would yield the distribution of density with depth.

Since the earth is known to possess a core, it cannot be homogeneous and there is no point in integrating Equations (21) and (25)

on the hypothesis that it is homogeneous. A further difficulty arises if the integration is carried out over the mantle. When this is done, it is found that the moment of inertia of the core is greater than $0.4\ M\ R^2$. A homogeneous mantle would thus require a core with a mass concentrated toward the outside. Since the core is a fluid, it is more probable to assume that the fluid core is more nearly chemically homogeneous than the mantle. Bullen (1953) considered these problems in detail and required that the density increase between 200 and 1000 km. Birch (1952) shows that between about 200 and 900 km, the rate of rise of velocity is too great for a homogeneous layer and indicates a gradual change of composition, or of phase, or of both. New phases are needed to account for the high elasticity of the deeper part of the mantle (below 900 km) and, because of this, Birch suggests a gradual shift toward higher pressure forms. The details of the possible phase transitions will be considered later.

The density distribution shown in Figure 2 has been derived using the velocities determined by Gutenberg (1959). An important point is that the moment of inertia derived from satellite and astronomical observations determines the magnitude of the density jump at the core-mantle boundary. This point appears to have been misunderstood by James and Kopal (1963).

The Earth's Free Oscillations

Following the great Chilean earthquake of May 22, 1960, the earth's free modes of vibration were observed for the first time. Their detection

broadened the spectrum over which a geophysicist may study the earth's interior. Prior to 1960, information regarding the earth's interior had been derived from detailed investigations of the arrival of the elastic body waves as recorded by seismographs. These waves travel different paths through the earth and contain most of the energy in the high-frequency part of the spectrum (10 - 0.1 cps). The use of low-frequency normal modes as a tool for interpreting the earth's interior is just beginning and a number of new results can be expected.

The notation adopted to describe the earth's free oscillations is similar in many ways to that used in spectroscopy; the similarity reflects the common mathematical structure of these two fields. The solution to the equations of motion of an elastic sphere can be separated into a function dependent upon radius and a function dependent upon the angular coordinates. The angular function is written as a sum of spherical harmonics

$$X_\ell^m = P_\ell^m (\sin \beta)\, e^{i(m\lambda - \omega t)} \quad , \tag{26}$$

where, P_ℓ^m is the associated Legendre function. The time-dependence has been included in the exponential where ω is the angular frequency of the oscillation (Alterman, et al., 1959). The exponential term describes a wave traveling with a speed of m/ω radians per second. If m is positive, the wave travels from west to east for a geographical coordinate system; if m is negative, it travels from east to west. The

two signs of m are symmetrical in a stationary sphere; the solutions for different m are degenerate. Rotation destroys this symmetry and creates important differences between waves traveling from west to east and waves traveling from east to west.

The numbers m and ℓ are familiar in quantum mechanics as the magnetic and azimuthal quantum numbers respectively; ℓ and m must assume integral values because of the conditions on the outer boundary and the conditions for regularity in the center of the sphere. The integral values determine the surface pattern of deformation associated with a particular free oscillation. The number of lines on the surface of vanishing displacement associated with the angular coordinate β is ℓ - m; the number of nodal lines associated with the angular coordinate λ is m. There will also be surfaces of zero particle displacement associated with the radial function. The description of the radial function is in terms of n, the radial node number.

The free oscillations can thus be characterized by three integers; ℓ and m determine the pattern of displacement on the spherical surface and n determines the number of internal nodal surfaces. The notation that has been adopted (MacDonald and Ness, 1961) is ${}_nS_\ell^m$, ${}_nT_\ell^m$ for the spheroidal and toroidal oscillations respectively. For example, in the ${}_0S_2$ oscillation, a sphere alternatively assumes a prolate and oblate form.

The eigen frequencies for an inhomogeneous earth can be determined by numerical computations (Alterman, et al., 1959; MacDonald and Ness,

1961). As an example, we first consider the toroidal oscillations which involve only the elastic mantle; the core-mantle interaction determines the boundary condition over the inner surface of the elastic shell. The effects of the inhomogeneous elasticity and the density on the oscillations are illustrated in Figures 3 and 4; the elasticity and density are those given in Figure 2. The energy levels shown in Figure 3 are normalized to provide a one-centimeter displacement on the surface (the observed displacement is much less; on the order of 10^{-6} cm). About 3×10^{21} erg are required to produce this displacement with a ${}_0T_2$ surface pattern in both the Gutenberg model earth and in a homogeneous earth having the average properties of the Gutenberg model. At higher frequencies it takes more energy to form the complicated surface pattern of displacement while maintaining this maximum surface amplitude. The needed energy is greater in the homogeneous model since the near-surface rigidity is greater than in the Gutenberg model.

The elastic energy in the oscillations is more or less evenly distributed over the entire mantle at low mode numbers (see Figure 5). At higher mode numbers, the elastic energy is concentrated toward the outer layers of the mantle. Thus, the high-order oscillations are used to describe the outer parts of the mantle, while the lower oscillations give information regarding the average properties of the mantle.

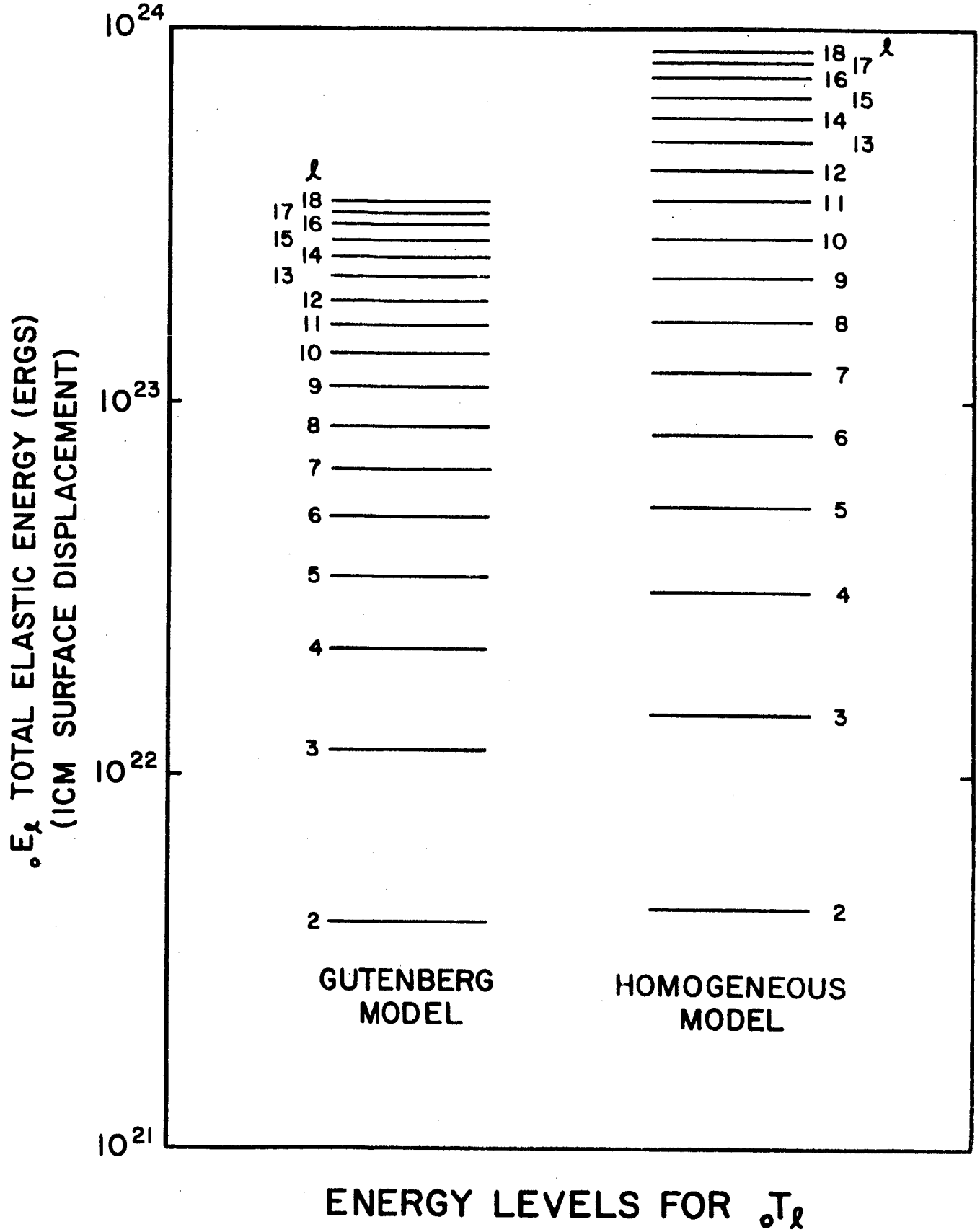


Figure 3: Total elastic energy in the toroidal oscillations. The energy is normalized to a one-centimeter surface displacement (MacDonald and Ness, 1961).

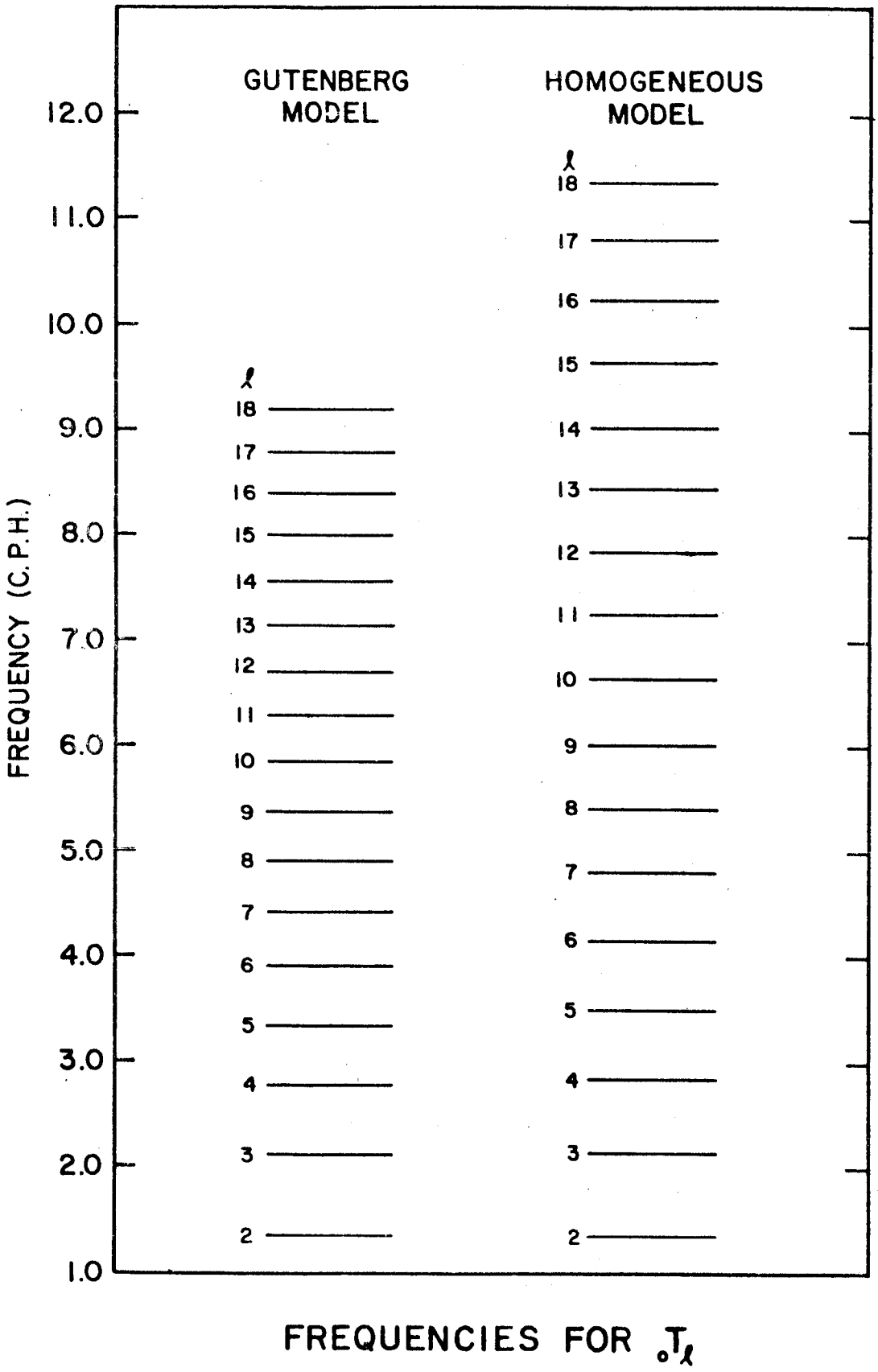


Figure 4: Comparison of the resonant frequencies for the toroidal oscillations in the Gutenberg and homogeneous models of the earth.

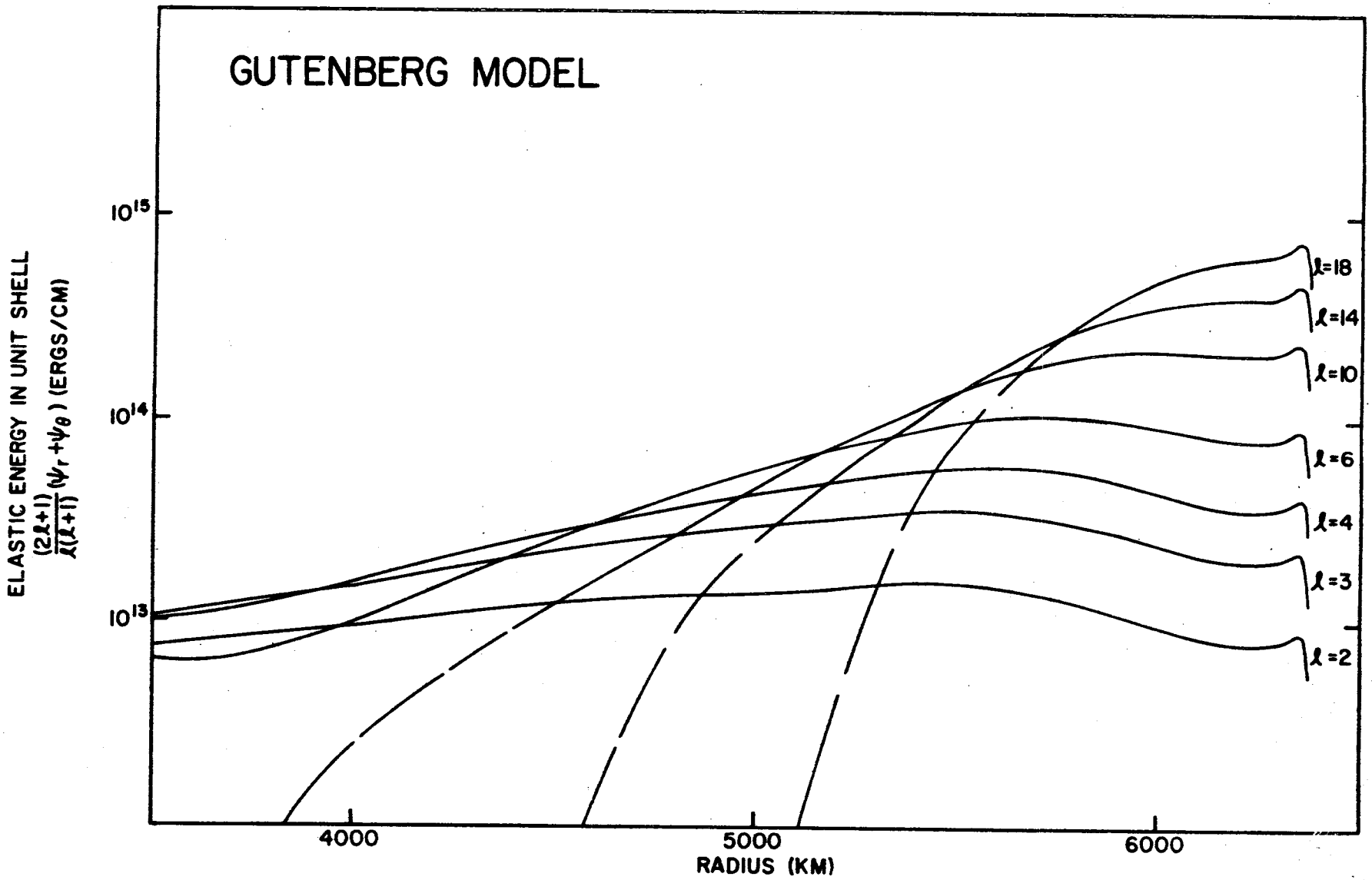


Figure 5: Elastic energy per unit radius in the toroidal oscillations for the Gutenberg model earth.

Observations of the Earth's Free Oscillations

Despite considerable theoretical efforts, only recently was an attempt made to observe the earth's oscillations. Benioff (Benioff, et al, 1959) constructed a strain-measuring seismometer in the form of a silicon glass rod 24 meters long with the purpose of investigating the low-frequency spectrum. The free oscillations excited by the Chilean earthquake were detected both on strain seismometers of the form initiated by Benioff and on sensitive gravimeters. The two types of instruments complement each other. The strain seismometer is sensitive to strain produced by both vertical and horizontal motion; it, therefore, records both spheroidal and toroidal oscillations. On the other hand, the gravimeter records only the vertical accelerations in spheroidal oscillations. A combination of the observations from the two instruments permits a separation and identification of the two types of modes of oscillations.

The power spectrum of the variations in gravity in Los Angeles for four days following the Chilean earthquake is shown in Figure 6. Figure 6 should be compared with Figure 7, a record of a quiet interval of 116 hours one month after the earthquake. This spectrum is almost structureless though there is a significant peak at 20.5 minutes.

The spectrum of the seismic disturbance is thus characterized by well-defined sharp peaks for periods between 1 hour and about 8 minutes. At higher frequencies, the isolated peaks begin to merge into a continuum as a result of the finite width of the individual peaks and their increased

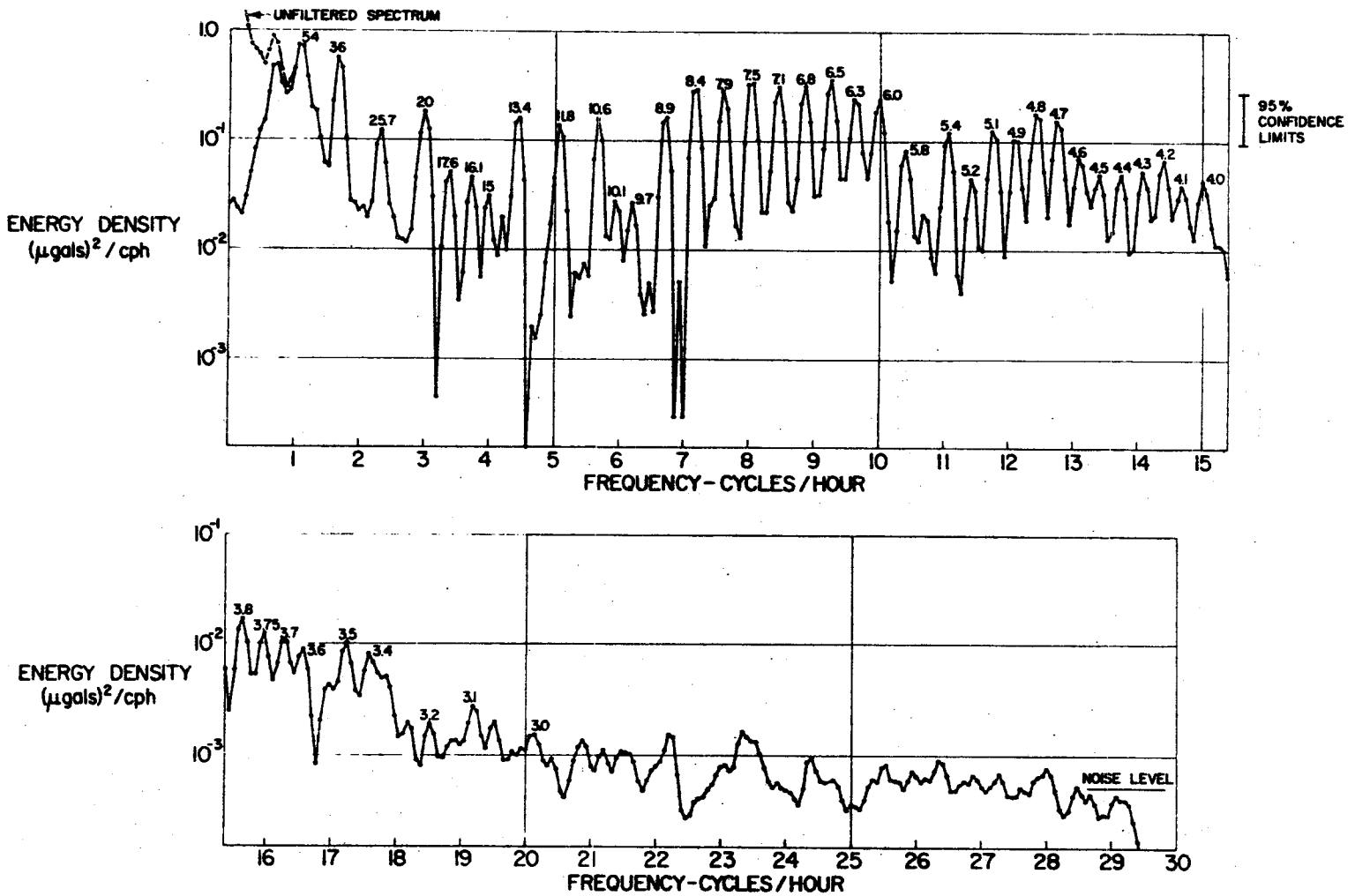

POWER SPECTRA ANALYSIS - UCLA EARTH TIDE GRAVIMETER
CHILEAN EARTHQUAKE - MAY 22, 1960 GMT
110 HOUR RECORD - 1 MINUTE INTERVAL - 0.1 μgal SENSITIVITY
0132 MAY 23 TO 1529 MAY 27 1960 (GMT)
UNFILTERED SPECTRUM
95% CONFIDENCE LIMITS
ENERGY DENSITY (μgals)² / cph
FREQUENCY - CYCLES / HOUR
NOISE LEVEL

Figure 6: Power spectrum of the gravity record after the Chilean earthquake of 22 May 1960. The University of California (Los Angeles) earthtide gravimeter; recording period, 23 - 27 May; 110-hour record; interval, one minute; sensitivity, 0.1 microgal.

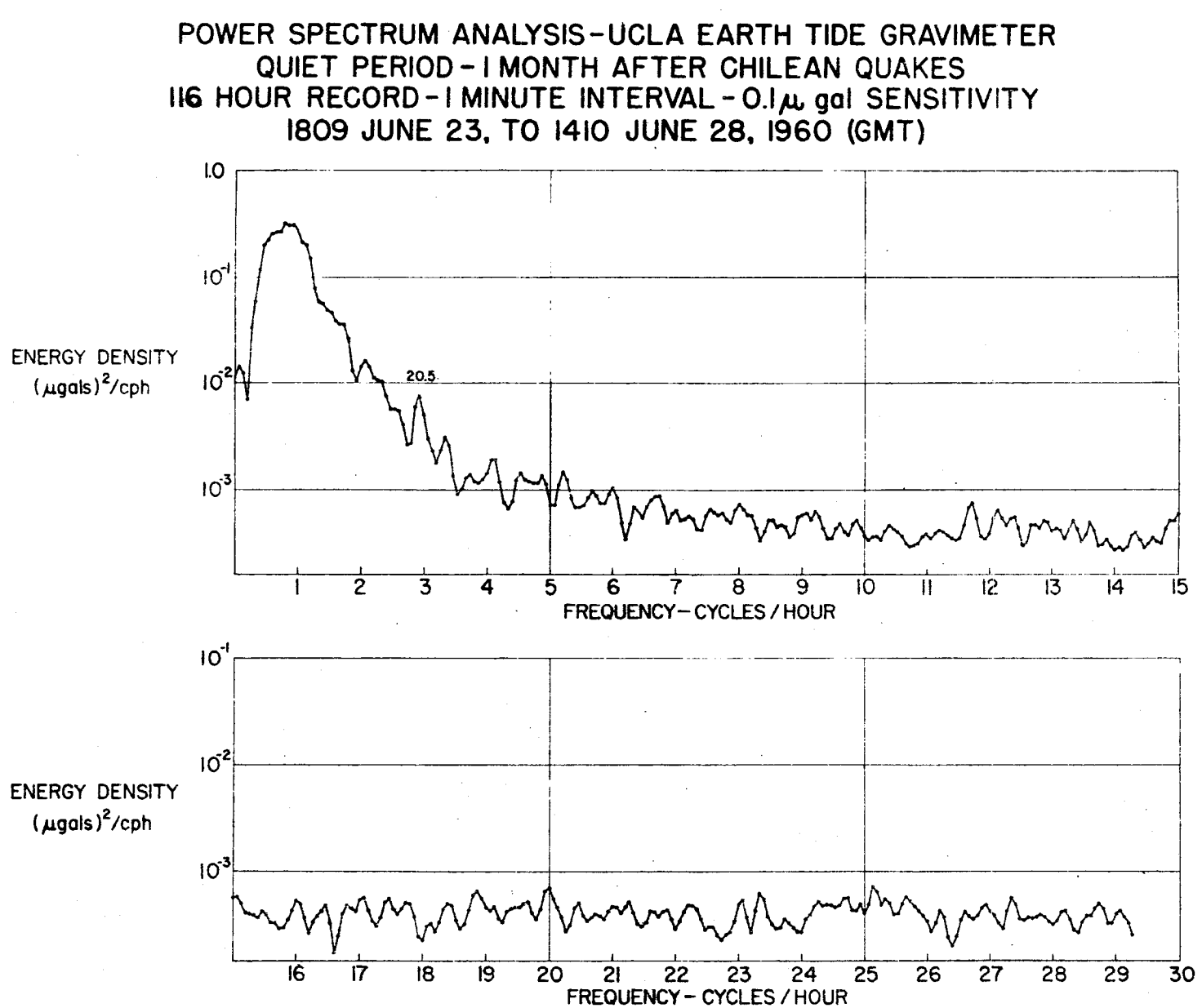

POWER SPECTRUM ANALYSIS-UCLA EARTH TIDE GRAVIMETER
QUIET PERIOD - 1 MONTH AFTER CHILEAN QUAKES
116 HOUR RECORD - 1 MINUTE INTERVAL - 0.1μ gal SENSITIVITY
1809 JUNE 23, TO 1410 JUNE 28, 1960 (GMT)
ENERGY DENSITY
(μgals)²/cph
1.0
10⁻¹
10⁻²
10⁻³
20.5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
FREQUENCY - CYCLES / HOUR
ENERGY DENSITY
(μgals)²/cph
10⁻¹
10⁻²
10⁻³
16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
FREQUENCY - CYCLES / HOUR

Figure 7: Power spectrum of a quiet period one month after the Chilean earthquake. University of California (Los Angeles) earthtide gravimeter; recording period, 23 - 28 June; 116-hour record; interval, one minute; sensitivity, 0.1 microgal.

number. A similar analysis carried out on strain records by Benioff, Press, and Smith (1961) and Alsop, Sutton, and Ewing (1961) gives similar results.

A comparison between the spheroidal modes calculated by Pekeris, Alterman, and Jarosch (1961b) and those measured by Ness, Harrison, and Slichter (1961) is shown in Figure 8. The Bullen model is similar to that of Gutenberg except that in the outer mantle the velocity increases monotonically rather than decreases as in the Gutenberg model. At low frequencies, the observations favor neither model since only the average properties of the earth are involved, but at higher frequencies the observations more closely fit the Gutenberg model.

Structure of the Spectrum

If the earth were a perfectly elastic, non-rotating sphere, then the spectral peaks would show up as individual lines broadened only by the data reduction techniques (instrumental broadening). Deviations from perfect elasticity and fluidity result in a natural broadening of the lines. The degree to which a given line is broadened and, alternatively, the rate at which a given peak decays provide a measure of the anelastic properties of the earth at the frequencies of the free oscillations.

From the width of the lines and the rate of energy dissipated per peak elastic energy, it is found that the Q is about 350 for the spheroidal oscillations. Thus, in the spheroidal oscillations, the

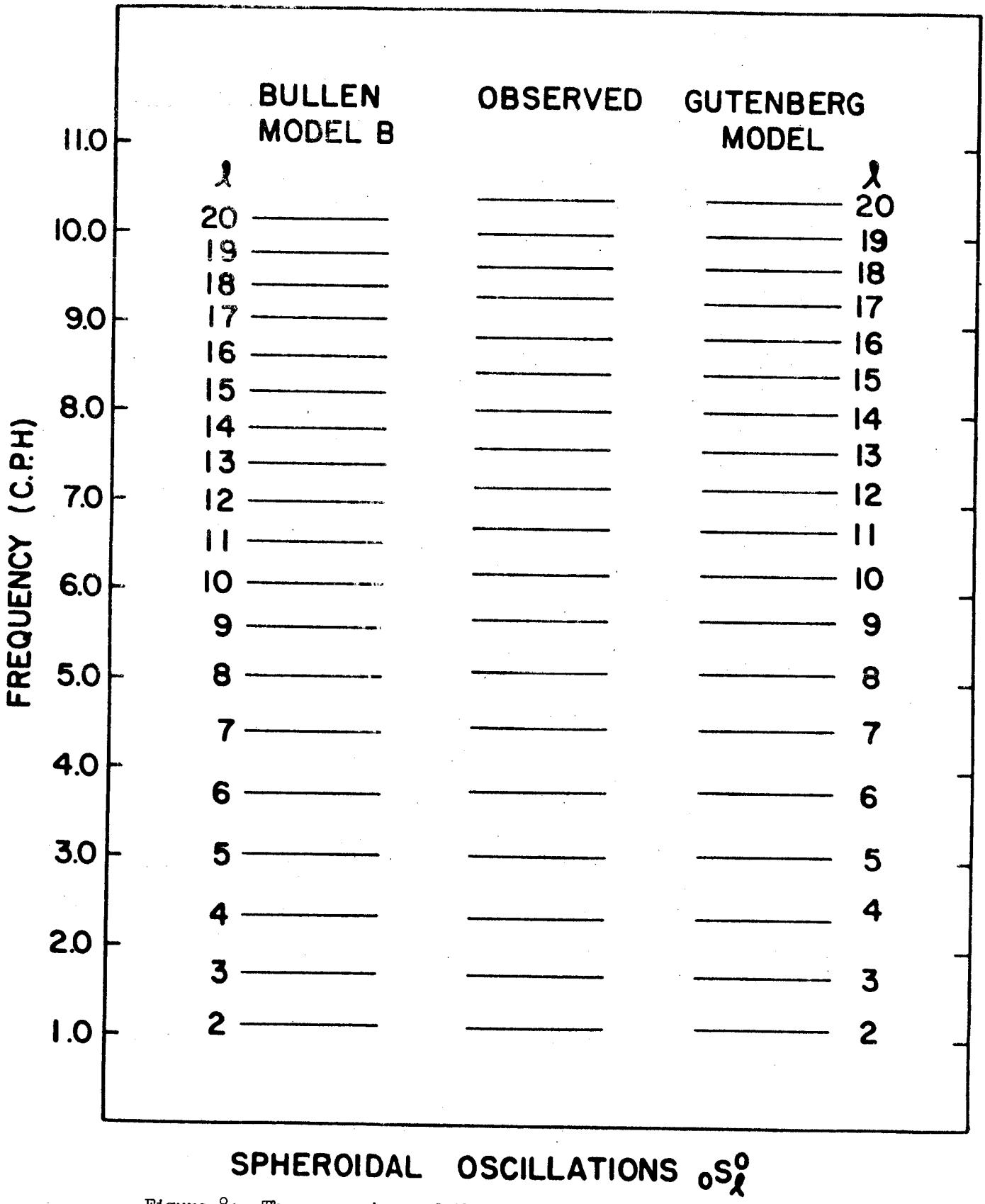


Figure 8: The comparison of the calculated and observed frequencies of the spheroidal oscillations.

earth rings as a rather poor bell. The estimate of the Q for toroidal oscillations is less certain and somewhat lower figures are indicated. The highest Q of all is shown by the radial oscillation. The indicated Q is greater than 5,000. This indicates that the deviations from elasticity in purely compressional motions are less than in shear.

There is additional time structure due to rotation. Rotation removes the degeneracy with respect to the mode number m. The effect of rotation on the oscillations can be qualitatively understood by recalling that a free oscillation is composed of a number of running waves. Waves traveling in the direction of the earth's rotation are carried forward relative to waves traveling in the opposite direction. The net effect is that the total pattern of surface deformation rotates relative to the earth. The removal of the degeneracy results in that any line is split into $2\ell+1$ lines. Detailed calculations of the splitting of the spheroidal oscillations have been made by Backus and Gilbert (1961), Pekeris, Alterman, and Jarosch (1961a), and MacDonald and Ness (1961).

Chemical Composition of the Earth

The chemical composition of the earth poses a problem both tantalizing and frustrating. Materials available at the surface can be sample-analyzed in great detail, but the main mass of the earth cannot and we must depend upon indirect approaches. It is usual to assume that the earth has a composition equal to that of chondritic meteorites. However, there are several serious objections to this hypothesis which we will consider later.

The major effort in classical geochemistry has been the accumulation of detailed analysis of rocks found at or near the earth's surface (Clarke and Washington, 1924; Rankama and Sahama, 1950; and Goldschmidt, 1954). From these analyses, the composition of the crust can be estimated. Table VII lists the crustal abundances.

Determining the composition of the remainder of the earth is extraordinarily difficult. We cannot make a direct chemical analysis of the material, but must depend on chance samples brought up from below. The interpretation of a given specimen, whether it is representative of deep material or the result of differentiation is not at all straightforward. Ross, Foster, and Meyers (1954) interpret the close mineralogical and chemical similarity of dunite and olivine bombs as indicating that these materials are the principal constituents of the upper mantle. However, as we shall show later, they cannot be if the thermal surface heat flow is to be accounted for.

Eclogite inclusions in kimberlite pipes are associated with inclusions of peridotite and dunite. They contain occasional diamonds and are enclosed within a diamond-bearing kimberlite. Because the diamonds must have formed at pressures corresponding to depth on the order of 80 to 100 km, the kimberlite inclusions probably represent material that crystallized at a depth of 100 km or more. Table VIII lists the atomic abundances of the major elements in eclogite, ultrabasic rock, peridotites, and dunites. A comparison of Table VII and Table VIII indicates the extent to which the lighter elements have

TABLE VII

Crustal Abundances
(after MacDonald, 1959a)

Element	Number of atoms per 10^6 atoms of silicon	Mass (g)
Na	1.21×10^5	6.26×10^{23}
Mg	1.26×10^5	6.63×10^{23}
Al	3.34×10^5	1.97×10^{24}
Si	1.00×10^6	6.28×10^{24}
P	4.68×10^3	3.16×10^{22}
K	5.70×10^4	5.11×10^{23}
Ca	1.24×10^5	1.07×10^{24}
Ti	1.11×10^4	1.62×10^{23}
Mn	2.60×10^3	3.02×10^{22}
Fe	1.19×10^5	1.45×10^{24}

TABLE VIII

Possible Samples of the Upper Mantle
(atomic abundances in number of atoms per 10^6 atoms of silicon)

(after MacDonald, 1959a)

Element	Eclogite	Ultrabasic Rocks	Peridotite	Dunite
Na	5.82×10^4	3.05×10^4	2.78×10^4	2.39×10^3
Mg	4.44×10^5	7.09×10^5	1.25×10^6	1.70×10^6
Al	2.54×10^5	1.32×10^5	1.29×10^5	2.50×10^4
Si	1.00×10^6	1.00×10^6	1.00×10^6	1.00×10^6
P	1.46×10^3	4.80×10^3	1.92×10^3	5.22×10^2
K	1.49×10^4	1.58×10^4	7.10×10^3	6.29×10^2
Ca	2.21×10^5	2.36×10^5	8.70×10^4	1.85×10^4
Ti	1.36×10^4	7.76×10^3	1.71×10^3	3.71×10^2
Cr	1.36×10^3	4.68×10^3		
Mn	3.29×10^3	2.96×10^3	3.66×10^3	3.34×10^3
Fe	2.41×10^5	2.17×10^5	1.58×10^5	1.67×10^5
Ni	6.94×10^2	3.35×10^3		

segregated in forming the earth's crust. The rock types listed in Table VIII have elastic properties similar to those observed seismically in the upper mantle. A further limitation on the chemical composition of the earth is derived from observations of the surface heat flow. These limitations are taken up in the next section.

Thermal Constitution of the Earth

The temperature in mines and wells exceeds the mean annual temperature at the surface and increases with depth. The rate of increase varies greatly from place to place, but the range is about 10 - 40 ^{o}C km^{-1}. A downward decrease in temperature implies an outward flow of heat. The outward heat flow is determined by multiplying the observed temperature gradient by the thermal conductivity of rocks. The mean heat flow is about 64 erg cm^{-2} sec^{-1}. The rate of flow of heat from the interior is thus about 20,000 times less than that arriving from the sun when it is overhead. The temperature of the surface is, therefore, controlled by the sun and not by the heat-producing, long-lived, radioactive elements and, possibly, by the initial heat stored within the earth since the time of formation.

Despite advances on many fronts, the distribution of temperature with depth within the earth remains uncertain, as does the principal mechanism by which heat is transported. At low temperatures, a solid transports heat by ordinary thermal conduction. At higher temperatures, large amounts of heat may be transported by radiation in optically transparent silicates. However, the combined effects of ordinary conduction

plus radiation may be swamped by convection provided that the material making up the mantle can undergo mass motion.

In the present section we will consider the major factors which influence the internal distribution of temperature. It is useful to base the discussion in terms of the problem of the development of the temperature distribution from a set of initial conditions. In addition to the initial conditions, we require the distribution of heat sources within the earth and the radial variations of such parameters as thermal conductivity, heat capacity, and density, provided that the heat is transported solely by conduction and radiation. The calculations then predict the present-day distribution of temperature and surface heat flux. These predictions can then be compared with the observations. In addition to the requirement of the observed surface heat flow, there is also the fact that the outer mantle of the earth is solid. The present distribution of temperature must, therefore, lie below the melting point of the matter making up the mantle. The melting points of multi-component silicate systems at high temperatures and pressures are still unknown but estimates for simple systems can be made.

Radioactive Composition of the Earth

Radioactive isotopes that have produced significant amounts of heat through the earth's history are distinguished by two characteristics:

1) The product of the isotopic abundance and the rate of heat generation is relatively large.

2) The half-life of the isotope is of the same order as the age of the earth.

The known isotopes that meet the two requirements are U^{238}, U^{235}, Th^{232}, and K^{40}. Isotopes with shorter half-lives may have played an important role in the initial stages of the development of the earth. Table IX lists the half-life and heat generation of the important isotopes.

The heat production in representative rock types as well as chondritic meteorites is listed in Table X. It can be seen that the heat production varies by about 3 orders of magnitude.

It is commonly assumed that the abundance of the radioactive elements within the earth equals that of chondritic meteorites (U = 1.1×10^{-8} g g^{-1}, Th = 4.0×10^{-8} g g^{-1}, K = 8×10^{-4} g g^{-1}). The principal argument for this assumption is the approximate equality of the surface heat flow and the rate of heat production in a chondritic earth (Birch, 1958). If the entire earth is made up of chondrites and if the heat currently being produced by the chondrites is reaching the surface, then the surface heat flow would average 59.4 erg cm^{-2} sec^{-1}. Provided only the mantle were made up of chondritic materials, then the surface heat flow would be 40.1 erg cm^{-2} sec^{-1}. The observed average over the surface is about 64 erg cm^{-2} sec^{-1}. As has been shown by detailed studies by MacDonald (1959b), the effects of the initial temperature and higher rate of heat production in the past raise the current value of the heat flow above the rate of production, but nonetheless, the agreement between the chondritic earth and mantle and the observed heat flow is striking.

TABLE IX

Half-life and Heat Generation of Important Heat Producing Isotopes
(see Aldrich and Wetherill, 1958; MacDonald, 1959b)

Isotope	Half-life (units of 10^9 years)	Heat Generation (ergs $g^{-1}sec^{-1}$)	Percentage Abundance of Isotope
U^{238}	4.51	0.94	99.27
U^{235}	0.71	5.7	0.72
Th^{232}	13.9	0.26	100.
K^{40}	1.25	0.288	0.0119

TABLE X

Heat Production in Rocks and Chondritic Meteorites
(after Tilton and Reed, 1963; MacDonald, 1959b)

Rock	Concentration in parts per 10^6			Heat Production (erg $g^{-1}year^{-1}$)
	Uranium	Thorium	Potassium	
Granite	4.0	14.8	35 000	285
Intermédiate	2.0	7.4	18 000	143
Basalt	0.8	3.0	7 500	58
Eclogite	0.043	0.16	530	3.2
Peridotite	0.006	0.022	10	0.38
Dunite	0.001	0.004	10	0.08
Chondrite	0.011	0.041	800	1.56

The difficulties with a chondritic earth or mantle have been discussed by Gast (1960) and Tilton and Reed (1963). Crustal rocks and rocks that might possibly originate in the mantle (eclogite, peridotite) show a ratio of potassium to uranium of about 10^4 (see Table X). The ratio in chondrites is 8×10^4. There are two possibilities:

1) The crustal and upper mantle material has been enriched with uranium relative to potassium.

2) The radioactive composition of the earth differs greatly from that of chondrites.

If this last possibility is correct, then there is little guide to the total radioactivity of the earth. The potassium content cannot be arbitrarily reduced while the uranium is fixed. An earth having the uranium abundance of chondrites with the ratio of potassium to uranium of 1×10^4 cannot produce the heat required to account for the present surface heat flow. A possibility is that uranium and thorium abundances are twice and the potassium is one-fourth that of chondrites ($U = 2.2 \times 10^{-8}$ g g^{-1}, $K = 2 \times 10^{-4}$ g g^{-1}). Such a composition would account for the present heat flow. In addition, a higher uranium abundance would be in closer agreement with current theories of nuclear synthesis (Hoyle and Fowler, 1963).

While the total amount of heat producers within the earth is limited by the observations of the surface heat flow, the ratio of short-lived potassium to long-lived uranium and thorium is important in considering the early stages of earth history. The variation of the rate of heat production for chondritic matter as a function of

time is shown in Figure 9. The heat production 4.5 Aeons ago was somewhat in excess of 8 times the present rate of heat production. If, on the other hand, the potassium ratio were 10^{-4} and the uranium content were 2×10^{-8} g g^{-1}, then the heat production 4.5 Aeons ago would only be about 4 times the present rate.

Surface Heat Flow Measurements

The most important quantity in any discussion of the thermal character of the earth is the amount of heat that is presently escaping from the earth's interior. The outward flux of heat by conduction per unit area per unit time is equal to the product of the thermal conductivity and the temperature gradient. The determination of heat flow requires separate measurements of thermal conductivity and of temperature gradient.

Systematic studies of thermal conductivities and temperature gradients have been carried out only since 1939. Birch (1954) reviewed measurements on heat flow and discussed about 40 determinations in continental areas and 25 at sea. Since then, a large number of measurements have been made. Lee (1963) assembled the presently available heat flow data and analyzed 614 values. Lee and MacDonald (1963) have carried out a detailed investigation of the heat flow using 757 observations. Of these, 92 were in continental areas and 665 in oceans.

The arithmetic means of the values are shown in Table XI. Lee and MacDonald (1963) also expanded the observed values in a set of

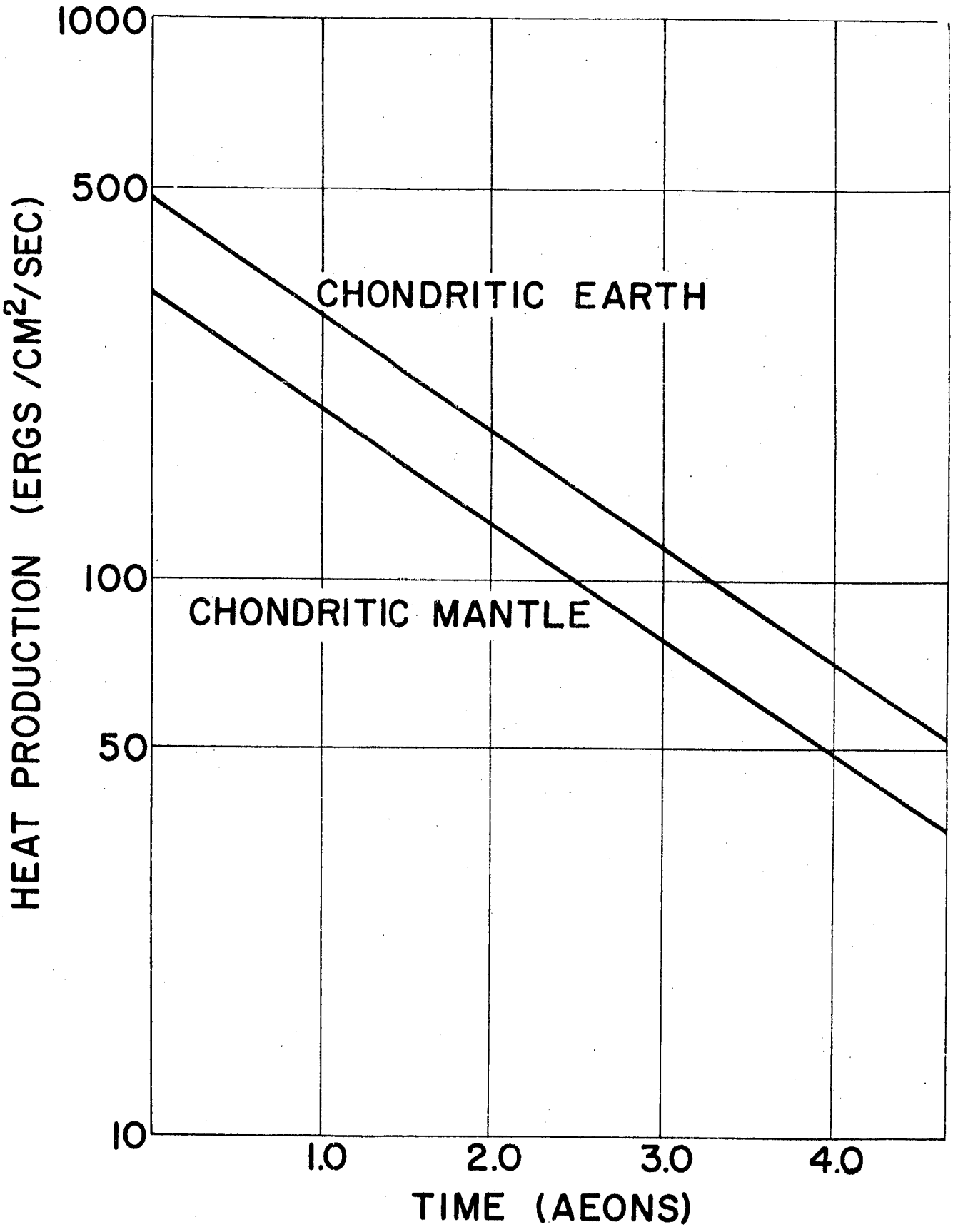


Figure 9: The heat production per column of unit surface cross-section. In the upper curve the total radioactivity of the earth is that of chondritic meteorites. The lower curve corresponds to an earth containing a chondritic mantle.

TABLE XI

Mean Heat Flow Based on 757 Observations
(after Lee and MacDonald, 1963)
in erg $cm^{-2}sec^{-1}$.

	Earth	Continent	Ocean
Arithmetic Means	67.4 (±1.8)	61.9 (±2.6)	68.2 (±2.0)
Orthogonal Expansion	63.9 (±3.4)	68.9	62.0

functions orthogonal to the observing stations. The resulting mean values are also listed in Table XI. The striking result is that the heat flow in continental and oceanic areas is very nearly the same. The global variation of surface heat flow is illustrated in Figure 10.

The heat flow is determined by the combined measurements of thermal conductivity and temperature gradient. The measurements will thus take into account the contribution of the transport of heat by the penetrative convection of magma. The measurements do not include the contribution to the escape of heat by lava deposited on the surface. Estimates of the rate at which lava is presently deposited on the surface vary considerably. Kuenen (1950) suggests a value of 1 km^3 per year. The total heat loss by one gram of lava crystallizing and cooling from 1000° to $0^{\circ}C$ is about 1.67×10^{10} erg g^{-1}. Of this, 4.2×10^{9} erg g^{-1} comes from the latent heat, 1.25×10^{10} erg g^{-1} comes from the specific heat. The average contribution of a cubic kilometer of lava to the surface heat flow is only 3.1×10^{-1} erg cm^{-2} sec^{-1}. The contribution of lava poured on the surface to the total heat flow is thus completely negligible if the estimate of a cubic kilometer per year is relevant. An alternative estimate follows from assuming that the total mass of the crust has at one time or another arrived at the surface as fresh lava. For the crustal mass of 2.45×10^{25} g produced in a time of 4.5×10^{9} years, the average rate of addition is 5.4×10^{15} g $year^{-1}$ as compared with the estimate of 3×10^{15} g $year^{-1}$ discussed by Kuenen. While the global contribution of volcanic activity is negligible, local volcanic activity can dominate and lead to an abnormally high heat flow

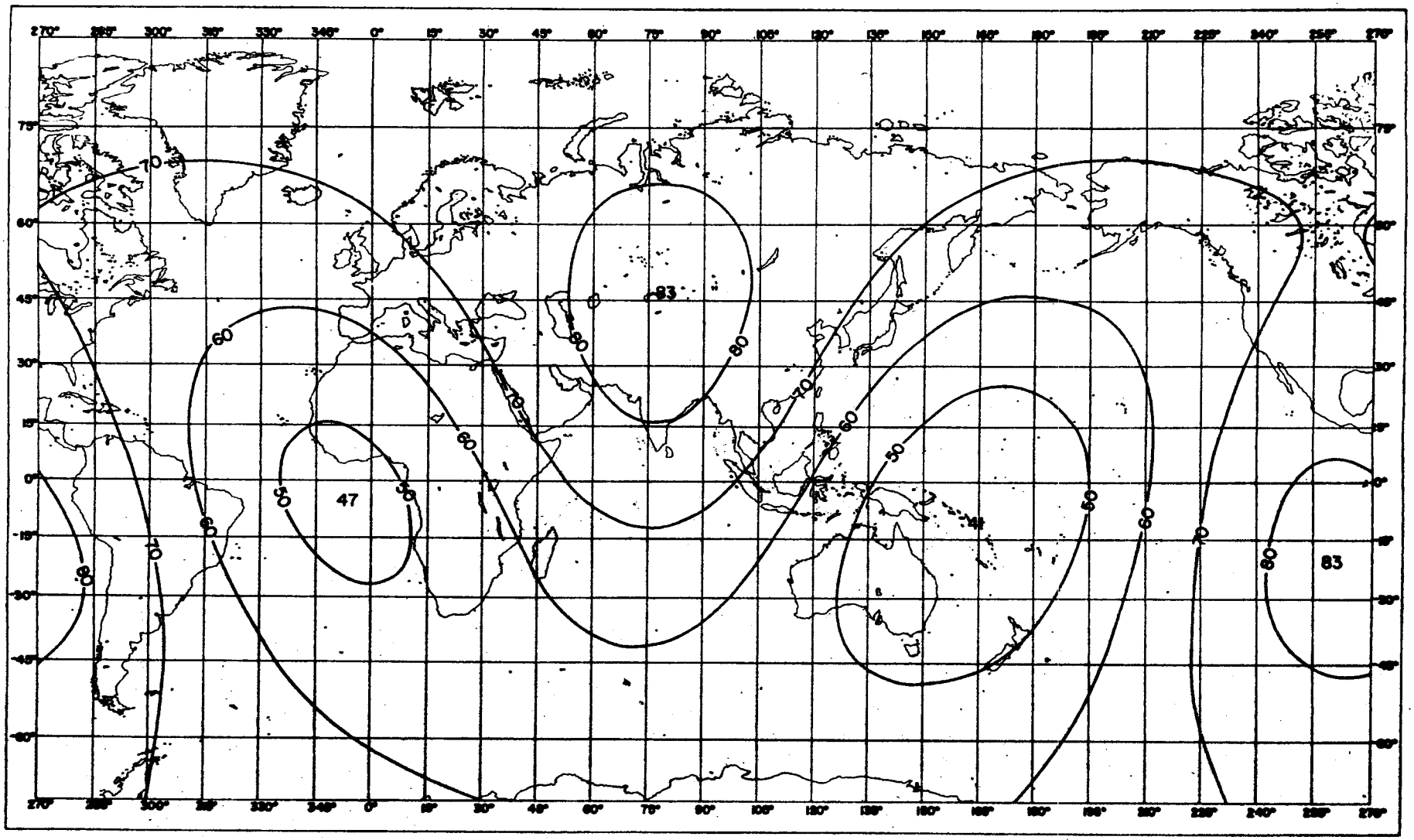


Figure 10: Orthogonal function representation of 757 heat flow values (Lee and MacDonald, 1963). Contours are in units of erg cm^{-2} sec^{-1}.

in the vicinity. In heat flow studies volcanic regions are avoided in continental areas but the sampling of the ocean floor is more nearly random.

Heat Flow and the Deep Distribution of Radioactivity

The structure of the continental and oceanic crust differs (see Table VI). The volcanic rock in the oceanic crust is usually considered to be basalt. The dominant material of the continental crust is a rock of intermediate composition. From Table X we note that the rate of heat production in the material above 40 km must differ greatly between continents and oceans. If the heat flows through the continents and ocean bottoms are approximately equal and if the heat productions of the surface rocks in the two regions are different, then there must be differences in the vertical distribution of radioactive material (Bullard, Maxwell, and Revelle, 1956; Bullard, 1954). Let Q_o represent the surface heat flow. The equilibrium heat flow at depth z is then $Q_z = Q_o - Az$ where A is the heat production per unit volume. For the heat production shown in Table X, it is seen that in the continental crust between 30 and 60 erg cm^{-2} sec^{-1} are produced and the heat flow in the continental areas is 68.9 erg cm^{-2} sec^{-1} (Table XI). The heat flowing from below a depth of 30 - 40 km into the crust is then between 9 and 39 erg cm^{-2} sec^{-1}. There are large uncertainties associated with the difficulties in correctly assessing the relative contributions of various rock types to the continental structure. If the volcanic material under the oceans is basalt and the sediments have the radioactivity of intermediate rocks, then the heat produced above 30 km is

between 3 and 6 erg cm^{-2} sec^{-1}. The heat flow in oceanic regions is 62 erg cm^{-2} sec^{-1} so that between 56 and 59 erg cm^{-2} sec^{-1} must be flowing from the depths into the crust. The difference in heat flow between oceans and continents at a depth of 35 km is then between 17 and 50 erg cm^{-2} sec^{-1}. The most probable figure is around 40 erg cm^{-2} sec^{-1}.

The low radioactivity of possible constituents of the upper mantle presents major problems in accounting for the heat flow under oceans (MacDonald, 1959a). If the upper 1000 km of the mantle are peridotite, then the heat produced in such a column is only 3.5 erg cm^{-2} sec^{-1} so that adding the production in the oceanic crust, the total heat flow would be about 8 erg cm^{-2} sec^{-1}. If the upper 1000 km has the radioactivity observed in eclogites, then the production under oceans is about 35 erg cm^{-2} sec^{-1}, or little more than half the required heat flow. If eclogitic radioactivity is assumed, then the radioactivity must be buried to at least 2000 km and some means for transporting the heat through these 2000 km must be imagined.

Thermal Conductivity of Silicate Material

Birch and Clark (1940) showed that the thermal conductivity of a wide variety of rocks shows surprisingly small variation with composition. Recent measurements at room temperature further establish the narrow variation of conductivity with composition (Clark and Niblett, 1956; Birch, 1950). The temperature dependence of conductivity has not been thoroughly explored experimentally. Birch and Clark measured

conductivities up to 400°C and found that the conductivities of poor conductors such as feldspar and quartz increase with temperature whereas the conductivities of most materials decrease slightly as the temperature increases. No experimental studies have been carried out on the variation of conductivity with pressure.

At high temperatures energy may be transferred within the solid by radiation rather than by lattice vibrations. Clark (1957) shows that the contribution of radiation to thermal conductivity is approximated by

$$K_R = 16n^2 s \frac{T^3}{3e} \quad . \tag{27}$$

In this expression, n is the effective index of refraction of the material, s is the Stephan-Boltzman constant, e is the sum of the absorption and scattering coefficients averaged over all wavelengths, and T is the temperature in degrees Kelvin. The striking feature of the contribution of radiation to conductivity is a strong temperature dependence. The higher the temperature, the greater will be the energy transferred by radiation provided that the variation of index in refraction and opacity do not overwhelm the T^3 dependence.

The opacity, e, and the index of refraction are determined by the properties of the material and, in general, will be functions of temperature and pressure. The variation of the index of refraction of a given material with pressure has not been investigated experimentally.

An estimate can be obtained by noting that for materials of different composition the index varies linearly with density. If the law of corresponding states holds, then the index should vary linearly with pressure. Within the mantle, the density varies from about 3.3 to 5.7 so that over the mantle of the earth a variation by a factor of 2 would be expected, with the material at depth having a higher index of refraction than that near the surface.

The rate at which radiation is transferred through a solid depends on the frequency of the radiation. Various mechanisms are known by which solids can absorb radiation and these become important at different frequencies. Absorption due to excitation of lattice vibrations by the radiation is strong in the infrared. This lattice absorption is relatively unimportant, since at high temperatures the energy density is low at these long wavelengths. Intrinsic absorption is due to the excitation of valence electrons to the conduction band across the fundamental energy gap. Intrinsic absorption is important in the ultraviolet for wavelengths less than about 0.4 microns. The transparency of silicates to radiation is thus limited at long wavelengths by the infrared absorption due to the lattice vibrations and at short wavelengths by absorption in the ultraviolet due to the excitation of electrons to the conduction band.

The region of high transparency in silicates lies in the visible to near-infrared. This is not true for semiconductors such as silicon and germanium, which are opaque in the visible. Because of the lower energy gap, the absorption edge in these materials lies at much longer

wavelengths than in silicates. Within the visible and near-infrared region of transparency, two processes can lead to absorption. Characteristic absorption peaks are associated with the presence of transition elements. The energy levels of the unfilled electron shells are split by the crystalline field and certain transitions between these split levels are allowed. These transitions lead to characteristic absorption bands of the transition elements. In silicates the most important transition element is iron which has a strong absorption peak about one micron. Titanium, manganese, and other elements will give rise to absorption bands.

Absorption between the peaks determines the contribution of the radiation to thermal conductivity. If there is one perfectly transparent region, then the material has an infinite thermal conductivity. The general level of absorption between absorption bands limits the energy transported by radiation. This general absorption is primarily due to free electrons. In the classical theory, free electrons will absorb at all wavelengths. The dependence of the opacity, ϵ, on the electrical conductivity is given by $60\pi\sigma/n$ where both the electrical conductivity, σ, and the index of refraction, n, vary with frequency. If we assume that the DC electrical conductivity is a sufficiently good approximation, the variation of opacity with temperature can be written as

$$\epsilon = \epsilon_o + \frac{60\pi\sigma_o}{n} e^{-E/kT} \quad , \tag{28}$$

where ϵ_o is the opacity at low temperatures and where the second term on the righthand side takes into account the temperature dependence of the conductivity, E is some characteristic excitation energy.

Most estimates of the contribution of radiation to the thermal conductivity are based on the determination of the absorption spectra rather than on actual measurements. In a few cases, the measurements of the spectrum can be made as functions of temperature and these indicate the closing off at high temperatures of the region of transparency for glasses containing transition elements. Clark (1957) measured the room-temperature absorption spectrum of several silicates in the visible and near-infrared. The radiative conductivity as calculated from the absorption spectrum is shown in Table XII where it should be noted that the ordinary conductivity at room-temperature is about 0.02 joules cm^{-1} sec^{-1} ^oC. The corresponding opacities are somewhat uncertain but a value of about 10 cm^{-1} accounts for the conductivity of diopside and garnets and a lower value is required for olivine.

Melting Relations at High Pressures

The fact that the mantle is solid could be used to limit the possible temperature distribution if the dependence of the melting point of the mantle material on pressure were known. As yet, there is little information on the variation of the melting temperature with pressure. Experiments that have been carried out refer to single-component silicates while the mantle material is undoubtedly

TABLE XII

Radiative Conductivity Calculated from Absorption Spectra

Mineral	K_R (joules $cm^{-1}sec^{-1}deg^{-1}$) 1000°K	1500°K	2000°K	2500°K
Olivine $(Mg,Fe)_2SiO_4$	0.297	0.860	1.45	2.02
Diopside $CaMgSi_2O_6$	0.067	0.238	0.44	0.725
Pyrope $Mg_3Al_2Si_3O_{12}$	0.004	0.021	0.075	0.184
Almandine $Fe_3Al_2Si_3O_{12}$	0.004	0.017	0.042	0.212

a multi-component system. The melting point of the single-component then provides some sort of upper limit. However, the melting curve of olivine, $(Mg,Fe)_2SiO_4$, which is the most likely constituent in the upper mantle is as yet unknown.

Boyd and England (1963) have determined the melting curves of diopside and albite. The measurements on diopside extend from 5 to 50 Kilobars and on albite from 10 to 32 kilobars. The melting curve of albite is intersected by the transition of albite to jadeite plus quartz at a pressure slightly in excess of 32 kilobars.

It is customary to represent the melting relations in terms of the Simon semi-empirical equation. The pressure and temperature along a fusion curve are given by

$$P = \frac{a}{B}\left[\left(\frac{T}{T_o}\right)^B - 1\right] \quad , \tag{29}$$

where a and B are empirical constants and T_o is the melting temperature at zero pressure. The initial slope of the melting curve can be used to determine the constant a and B is determined by the initial temperature. The coefficients in the Simon equation are given in Table XIII for diopside and albite.

If the diopside melting curve is extrapolated to the core-mantle boundary where the pressure is about 1.4×10^6 bars, the melting temperature is 3750°C.

TABLE XIII

Coefficient for Simon Equation
(after Boyd and England, 1963)

	T_o °K	Initial Slope deg/1000 bars	a	B
Albite	1391	14	19.5	5.1
Diopside	1665	15.4	23.3	4.64

Convection

Much of the current speculation regarding the transport of heat within the earth and the other terrestrial planets centers on the possibility of thermal convection (Kopal, 1963; Runcorn, 1963). In the earth, convection in the mantle is usually imagined as a regular, cellular motion extending hundreds of thousands of kilometers in depth (Vening-Meinesz, 1962). There is no direct evidence of this motion, but because of the lack of sound theoretical restrictions on the nature of the convection currents, models can be constructed to produce a great variety of results. This type of large-scale convection is not to be confused with penetrative convection where local regions undergo melting and rise through fractures in the overlying material. There can be no doubt that this motion occurs; the existence of large-scale organized convective systems within the earth remains debatable. Because of the wide applicability of the notions of convection to the internal structure of the moon and other planets, we consider the problem of convection within the earth in some detail.

Theoretical work on mantle convection assumes that the mantle can be approximated as an incompressible, homogeneous fluid of zero strength (Pekeris, 1935; Hales, 1936; Heiskanen and Vening-Meinesz, 1958). Such theoretical treatments of convection are subject to three major criticisms:

1) The mantle is assumed devoid of strength; for the long-term stress system, the material behaves as a fluid with a classical viscosity. In application, it is usual to mention

the finite strength and then ignore it. The evidence reviewed in the previous section clearly shows that the mantle behaves as a material with a very high viscosity (10^{26} c.g.s.) or has a strength on the order of 200 bars. Theoretical criteria for instability in a material with finite strength have not been discovered. Even supposing convection to occur, we are completely ignorant of the form it would take. Would it, for example, be a regular, cyclical motion or be initially highly irregular? It is clear that beyond the yield point the anelastic properties of the mantle cannot be described in terms of viscosity.

2) Current theories of convection describe the temperature difference and then calculate the motion in the homogeneous fluid. The situation in the mantle and in other planets would be the exact opposite. The heat sources would be distributed throughout the convecting system and the driving mechanism for the presumed convecting material would be tied to the material itself. There is no solution of a comparable problem even for a simple viscous fluid; in this case, it is not known how the motion reorganizes the heat sources. It may be that such a convection would tend to concentrate the heat sources so as to maximize the heat flux, but this is sheer speculation.

3) In the usual theory of convection in a viscous fluid, the controlling parameters, density, viscosity, conductivity, etc., are taken as constant. The large differences in density with depth and the implied variation of the parameters are neglected. Density fluctuations are assumed to be due only to variations in

temperature. This approximation is valid provided that the change in density due to pressure or other inhomogeneities is small compared to the variations in density resulting from temperature fluctuations. For a convective system extending over a depth of 1000 km, this requirement fails by a large margin.

The usual argument for convection is that even though the theory is not well understood, the observational evidence for convection is so strong that convection must exist. We now briefly review this evidence.

1) Strong gravity anomalies are associated with deep ocean trenches. Vening-Meinesz (1948) suggests that the deficiency in mass implied by these strips is due to the down-buckling of light crustal rocks into the denser material of the mantle. This down-fold is supposed to be the site of a down-going convective current pulling the crust inward. However, seismic and gravity measurements carried out by Worzel and Shurbet (1954) show that there is no crustal thickening as required by the down-fold hypothesis.

2) Vening-Meinesz (1962) has noted irregularities in the spherical harmonic representation of the earth's topography. Let a_n^m, b_n^m represent the coefficients in the spherical harmonic expansion of topography. Then the average of the surface harmonics over the sphere of surface σ is given by

$$\frac{1}{4\pi}\int_{\sigma} (Y_n)^2 \, dr \quad , \tag{30}$$

where

$$Y_n = \sum_{m=o}^{n} (a_n^m \cos m\lambda + b_n^m \sin m\lambda) \, P_n^m(\sin \beta) \quad . \tag{31}$$

In Vening-Meinesz's analysis, the surface harmonics for the topography show a peak in the values n=3 to n=5. This correlates with modes for a viscous fluid motion that would be thermally excited in a spherical shell having a 3000-km thickness. Vening-Meinesz's hypothesis is that the surface topography is an expression of the deeper structure and that the topography of the deeper structure is indeed due to currents; thus, surface harmonics representing gravity should also show a peak at the corresponding harmonics. The test of this hypothesis can be made by evaluating the mean surface harmonics for the non-hydrostatic part of the gravitational field. The data is yet limited, but the surface harmonics derived from Kaula (1963b) and Kozai (1962b) values are

n	Y_n
2	5.59×10^{-6}
3	2.07×10^{-6}
4	1.08×10^{-6}

The mean square of the surface harmonics drops off regularly with degree of the harmonic. Indeed, these values coupled with values

for the higher-order zonal harmonics suggest that the mean square of surface harmonic may drop off as $n^{-5/2}$. Thus the agreement between the harmonic representation of topography and that to be expected for the homogeneous shell of fluid does not hold for the gravity field.

3) The thickness of the sediments on the sea floor is also considered to be evidence for convection (Hess, 1962). The sedimentation since Pleistocene has been at the rate of about 1 centimeter per thousand years and the Mohole test drilling indicates that the rate may have held since the Miocene. The present thickness of the sediments on the ocean floor would then accumulate in something like 200 - 300 million years. Because of the short time-scale, Hess argues that the ocean floor must be renewed from time to time. Convection currents are presumed to provide the mechanism by which the ocean floor is periodically rejuvenated.

The argument from the rate of sedimentation is reminiscent of attempts to date the earth from the salt content in the sea. Indeed, at current rates of inflow, the age of the ocean derived from the salt now present in the sea is about 300 - 400 million years. Rather than requiring a clean sea floor once every 200 million years, the present rate of sedimentation and the observed thickness suggest, as does the salt content of the oceans, that the rate of erosion has varied widely in geologic times. There are good reasons for supposing that both the present and the late Tertiary have been times of considerable relief during which the rates of sedimentation were considerably above the long-term geologic average.

4) The paleomagnetic data discussed by Runcorn (1962a), Cox and Doell (1960), and others can be interpreted in terms of motions of the continents relative to each other. This interpretation presumes that the field has remained dipolar over geologic time. The theory of the causes of the earth's magnetism is far from being established and one can argue that the observed reversals in the field suggest a rather short time-scale for the field and intervals during which the internal field may not have been dipolar. Further, there is accumulating paleontological evidence that argues strongly against continental drift (Axelrod, 1963).

None of the arguments for convection appear particularly convincing; on the other hand, there are strong reasons for rejecting convection as it is usually presented. The first of these reasons is concerned with the thermal requirements of mantle convection. Let ΔT represent the temperature difference between the up-going and down-going currents of a convecting cell. In the steady state, the viscous forces equal the buoyancy forces and this requires

$$g\rho\alpha\,\Delta T = \frac{\mu v}{(R-r)^2} \quad . \tag{32}$$

where,

g = gravitational acceleration,

ρ = density.

α = volume coefficient of thermal expansion,

μ = viscosity,

v = velocity of the convection currents,

and where the variation of the parameters with depth is neglected. For mantle convection, the two uncertain quantities in the above expression are the velocity and the viscosity. It is conventional to adopt 10^{22} c.g.s. as the viscosity. This value refers to a questionable interpretation of the local rebound in glaciated regions. An alternative estimate of the viscosity can be obtained from the figure of the earth. The theoretical flattening of the earth is greater than the observed flattening. This has been interpreted by Munk and MacDonald (1960a,b) to be the result of the slowing down of the earth by tidal friction and a lag in the establishment of the equilibrium figure.

$$\Delta T = 2 \times 10^{10} v \quad . \qquad (33)$$

Runcorn (1962b) argues that if convection is to play any role in geology, velocities on the order of 1 - 10 cm year^{-1} are required. If the motions along fault planes are interpreted as due to convection, velocities on the order of 1 cm year^{-1} are required. On the convective theory the paleomagnetic evidence again suggests values on the order of centimeters per year. In order to account for a velocity of 1 cm year^{-1}, a temperature difference between the upgoing and down-going column would have to be 650°. The heat carried

by such a convective system is of the order of $\rho c_p \Delta T v$, where c_p is the specific heat. A temperature difference of 650° with a velocity of 1 cm $year^{-1}$ would require a heat transport by convection of 650 erg cm^{-2} sec^{-1}. This is an order of magnitude greater than that observed (see Table XI).

Taking realistic values of the viscosity, it is seen that the thermal requirements of convection in the mantle are so severe that the hypothesis can be dismissed. Further arguments against large-scale convection can be marshaled from the observed equality of heat flow over continents and oceans and the equality in mass per unit area. It is difficult to imagine large-scale horizontal motions in which these conditions are met. In addition, there is the seismic evidence that the materials under continents and under oceans differ. Large-scale horizontal convection would destroy such differences.

Calculation of Internal Thermal Conditions

The internal thermal state of the earth can be investigated by assuming an initial temperature distribution, a distribution of heat sources, and a distribution with depth of such parameters as the density, heat capacity, thermal conductivity, and opacity. The temperature at the outer surface is assumed to be constant with time. The problem is then to determine the temperature distribution at later times. The equation governing the development of the temperature field is

$$\frac{\partial T}{\partial t} = \frac{1}{\rho c_p} [\nabla \cdot (K \nabla T) + A] - \underset{\sim}{v} \cdot \nabla T \qquad (34)$$

where,

ρ = specific density (g cm^{-3}),

c_p = heat capacity at constant pressure (erg g^{-1} $^oC^{-1}$),

K = total thermal conductivity (erg cm^{-1} sec^{-1} $^oC^{-1}$),

A = rate of heat production (erg cm^{-3} sec^{-1}),

$\underset{\sim}{v}$ = velocity (cm sec^{-1}).

The term containing thermal conductivity is a measure of the heat flux in and out of the unit volume. A is the rate at which heat is produced within the unit volume by radioactivity and is a function both of the position coordinate and of time because of the slow decay of radioactive elements. If a radioactive element is producing heat at a rate dQ/dt today, it produced heat at a rate $e^{\lambda t}$ dQ/dt at a time T years ago where λ is the disintegration constant of the radioactive element. The term involving the velocity on the righthand side is the time rate of change of temperature resulting from the advection of heat into the unit volume. It is this last term which couples the temperature equation to the equations of motion. The problem is much simplified if this term can be neglected.

In order to evaluate the advective term, we calculate the temperature field neglecting $\underset{\sim}{v} \cdot \nabla T$ and then compare the relative magnitudes of $A/\rho c_p$ and $\underset{\sim}{v} \cdot \nabla T$. The maximum radial velocities are reached during the

early history, when the rate of heat production in a unit column cross-section far exceeds the surface heat flow. The maximum radial velocities are on the order of 100 km per 10^9 years. The corresponding temperature gradients over the upper part of the mantle are then 10^0 per kilometer, so that the convective term is on the order of 3×10^{-14} ^oC sec^{-1}. The rate of heat production for a typical sub-oceanic mantle during the first 10^9 years of earth history is of the order of 8.7×10^{-6} erg cm^{-3} sec^{-1}. The term $A/\rho c_p$ is then 2×10^{-13} ^oC sec^{-1}. Thus, during the early stages of earth history where advection of heat is greatest, the advective term is an order of magnitude less than the term due to internal heat production. At later stages, the maximum radial velocities are on the order of a kilometer per Aeon, with the result that the advective term is about 2×10^{-16} ^oC sec^{-1}. The heat production term for the same time is 3×10^{-4} ^oC sec^{-1}. It thus appears that over the entire earth history, the heat production term is one to two orders of magnitude larger than the advective term. Though the heat production term dominates the advective term, the predominance is not great enough to ensure that the coupling of the motion field to the thermal field is entirely negligible even in the case where convection is neglected.

In order to investigate the development of temperature within the earth and the effects of the differing vertical segregations of heat sources, it is sufficient to consider the problem in two spatial dimensions. Equation (34) then takes the form

$$\rho c_p \frac{\partial T}{\partial t}(r, \theta, t) = \frac{1}{r^2} \frac{\partial}{\partial r}\left[r^2 K(r, \theta) \frac{\partial T}{\partial r}(r, \theta, t) \right] \qquad (35)$$

$$+ \frac{1}{r^2} \sin\theta \frac{\partial}{\partial \theta}\left[\sin\theta \, K(r, \theta) \frac{\partial T}{\partial \theta}(r, \theta, t) \right] + A(r, \theta, t) \quad ,$$

where θ is the angular coordinate.

Lubimova (1958) and MacDonald (1959b) have carried out detailed calculations on the development of temperature within a spherically symmetrical earth. MacDonald (1963) has extended these calculations taking into account the differing distribution of heat sources under continents and oceans. Several general features of the spherically symmetrical calculations are important. The principal effect of radiative transfer of energy on the distribution of temperature is the flattening of the temperature gradient with depth. Because of the contribution from radiation, the temperature required to remove heat from the deep interior decreases as the conductivity increases with temperature. Near the surface the temperature gradient is dependent only on the lattice conductivity of the solids and the near-surface distribution of heat sources. For a wide variety of models of the earth, a near-surface gradient of 10 - 20 degrees per kilometer results. Since the temperature gradients are greatest near the surface, the melting temperatures are most closely approached in the outer few hundred kilometers of the earth.

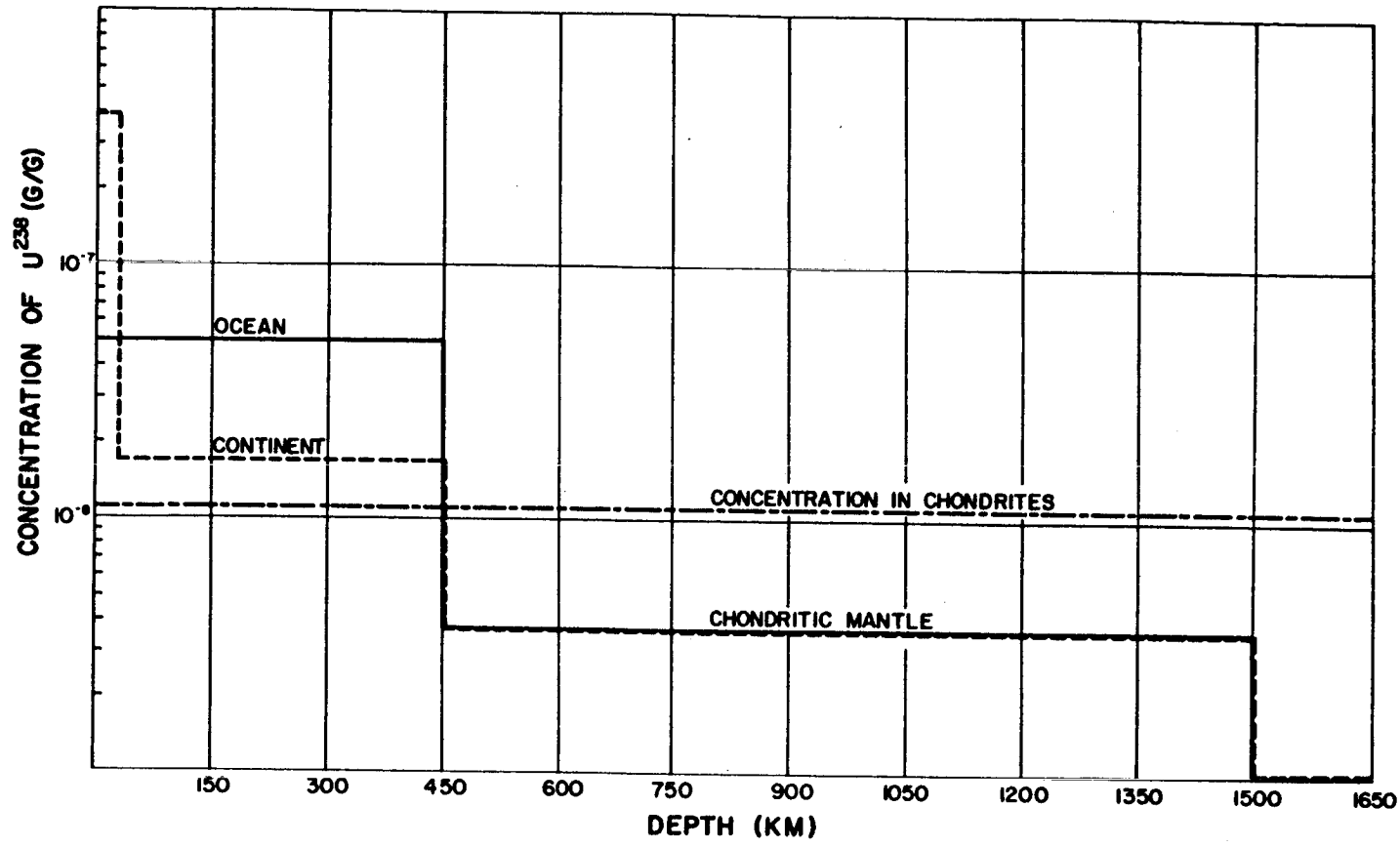
The present-day heat flow from a wide variety of earth models having a chondritic radioactive composition and very different initial

conditions and distribution of heat sources all lead to present-day heat flows nearly equal to the observed 64 erg cm^{-2} sec^{-1}.

As an illustration of the distribution of temperature with depth, we consider calculations carried out by MacDonald (1963). In these calculations, the total radioactivity of the earth was taken to be equal either to that of an earth made up of chondritic materials or an earth in which only the mantle had the radioactivity of chondrites and the core contained no radioactivity. The vertical distribution of heat sources under continents and oceans is shown in Figures 11 and 12 with the ratio of thorium to uranium taken as 3.7.

There are no known continents as old as the earth. However, a number of age measurements indicates that there are rocks having an age of about 3×10^9 years (Tilton and Davis, 1959). We consider the development of temperature in continents formed 3×10^9 years ago both for an initially low temperature given by Curve B in Figure 13 and for a higher initial temperature (Curve A of Figure 13).

The development of temperature with time for a continent formed 3×10^9 years ago at the low-temperature (Curve B of Figure 13) is shown in Figures 14 and 15. The radioactivity is taken equal to that of a chondritic mantle and the vertical distribution is as shown in Figures 11 and 12. The isotherms move inward with time. At present, the 1500° isotherm is at a depth of 190 km under the oceans and at 225 km under the continents.



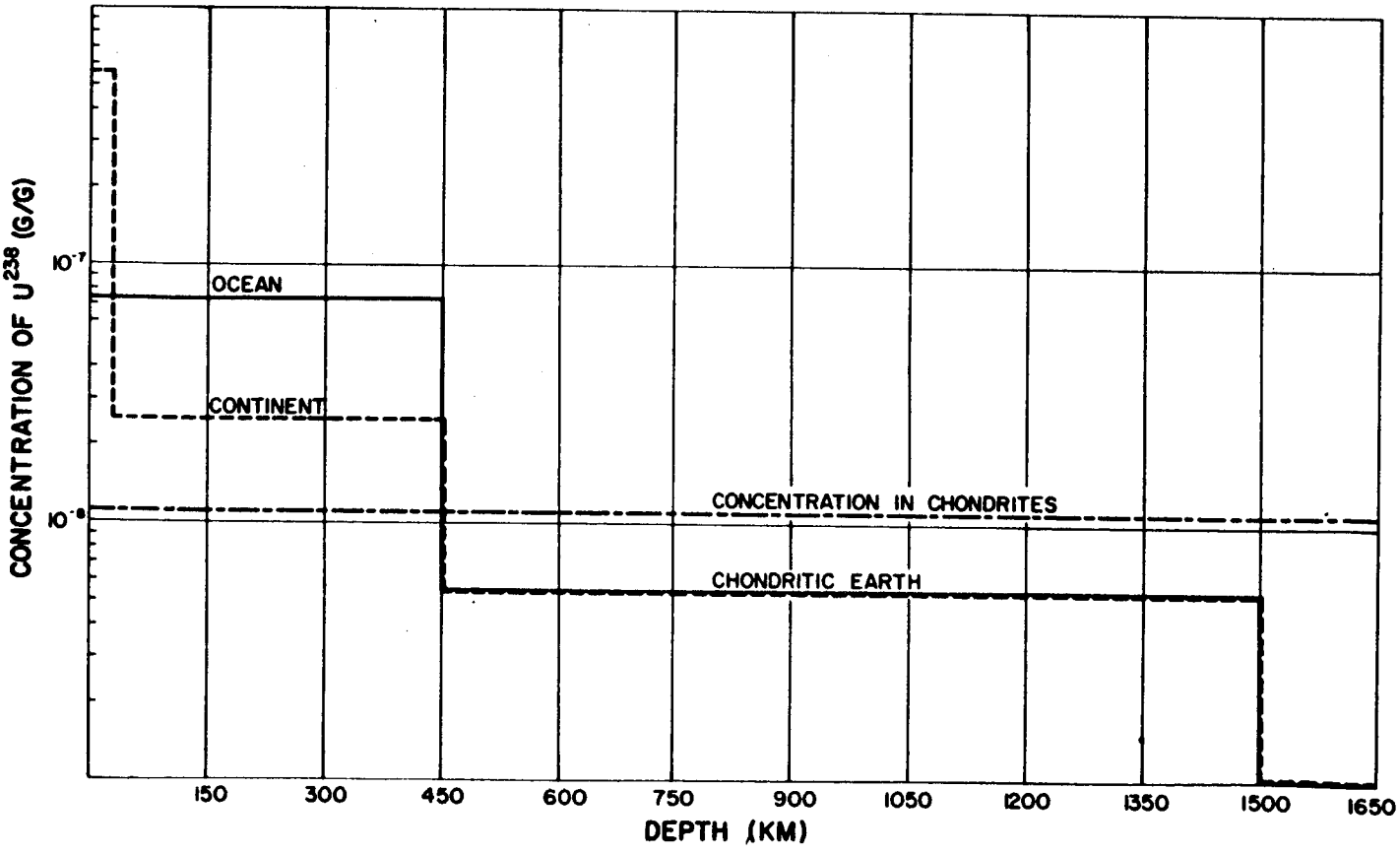
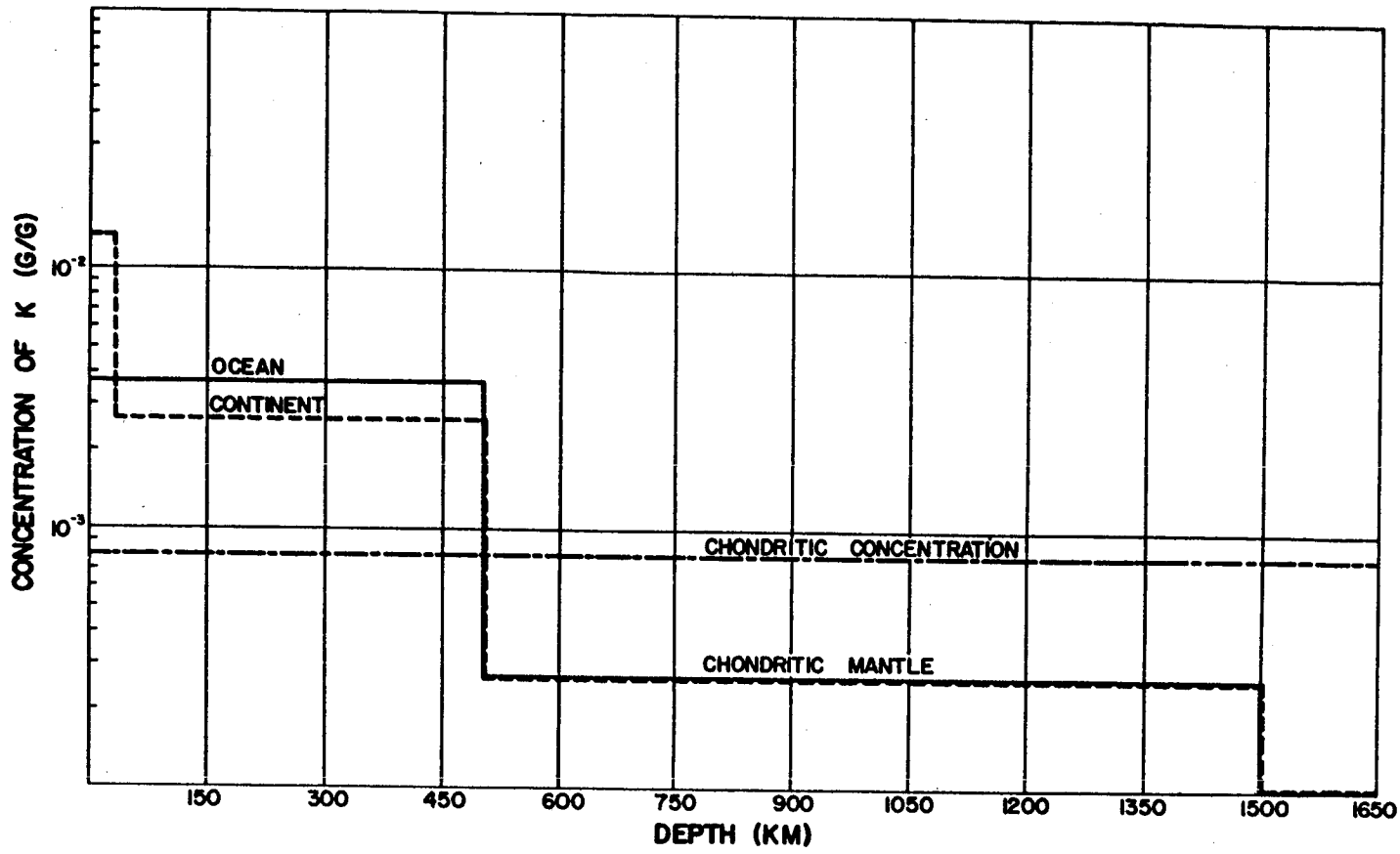


Figure 11: Vertical distribution of uranium in g g^{-1} for the models illustrated in Figures 14 - 23.



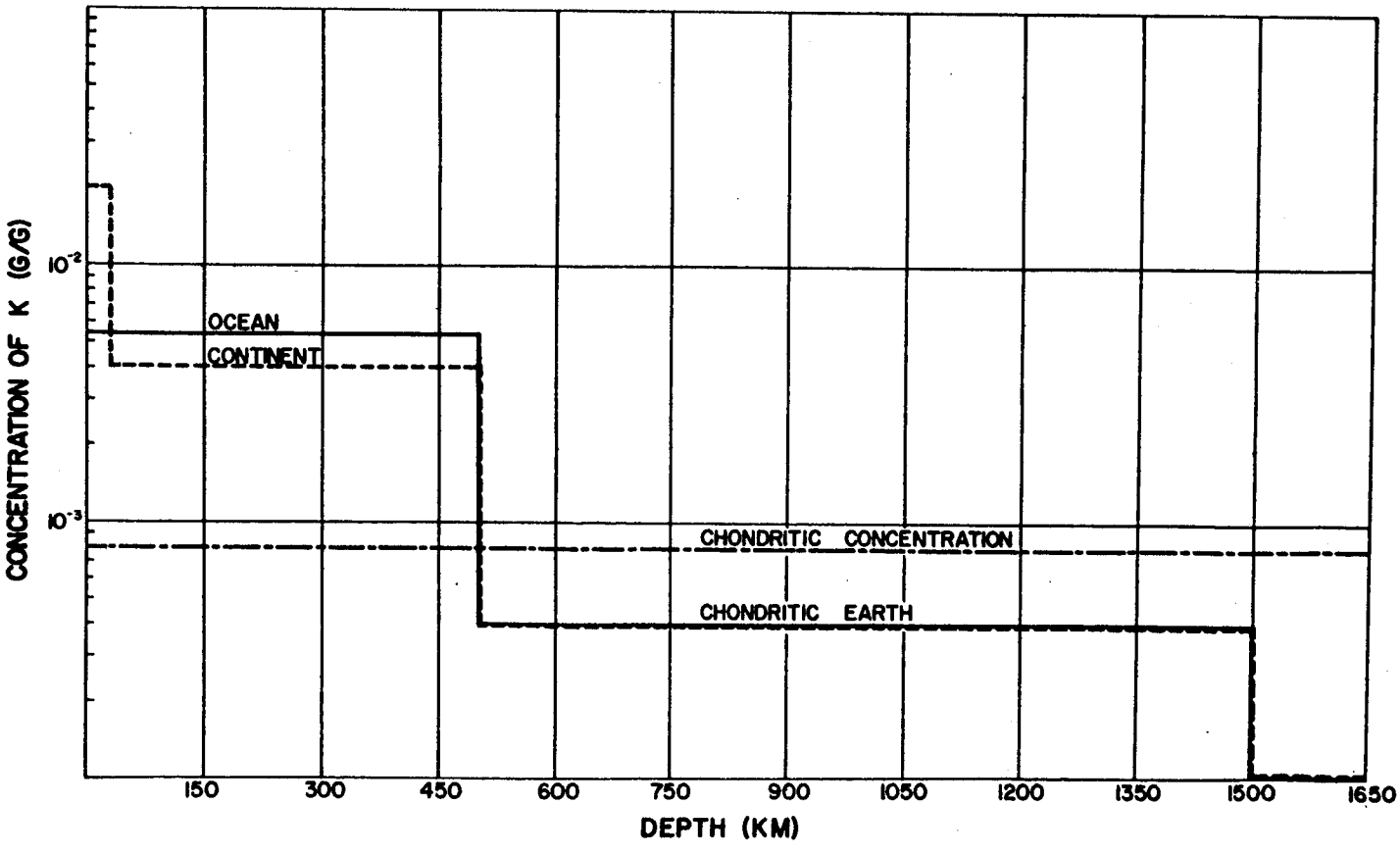


Figure 12: The concentration of potassium in g g^{-1} for the models illustrated in Figures 14 - 23.

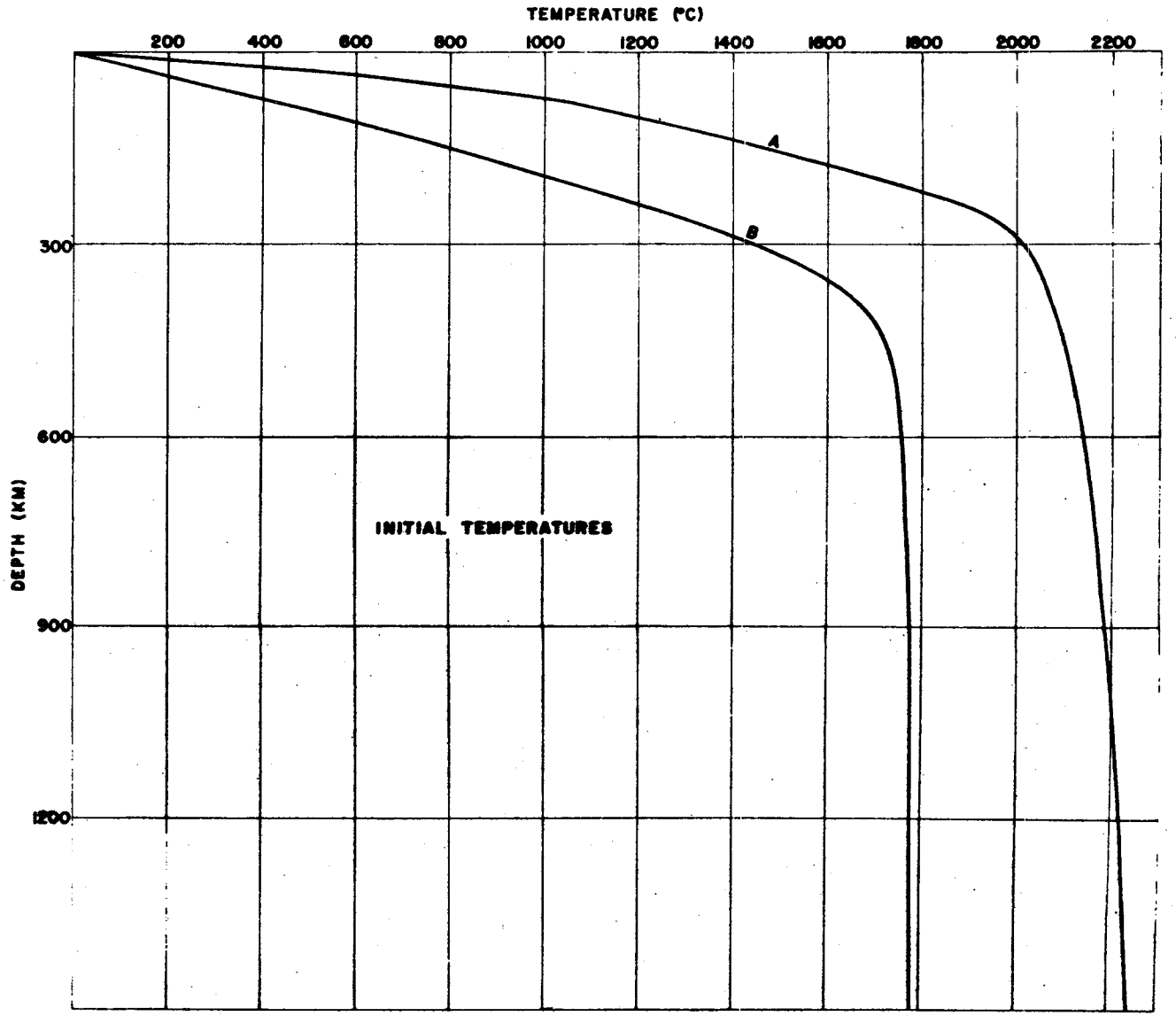


Figure 13: The initial vertical distribution of temperature used in the models illustrated in Figures 14 - 23. Curve A is referred to in the text as a high initial temperature; Curve B, a low initial temperature.

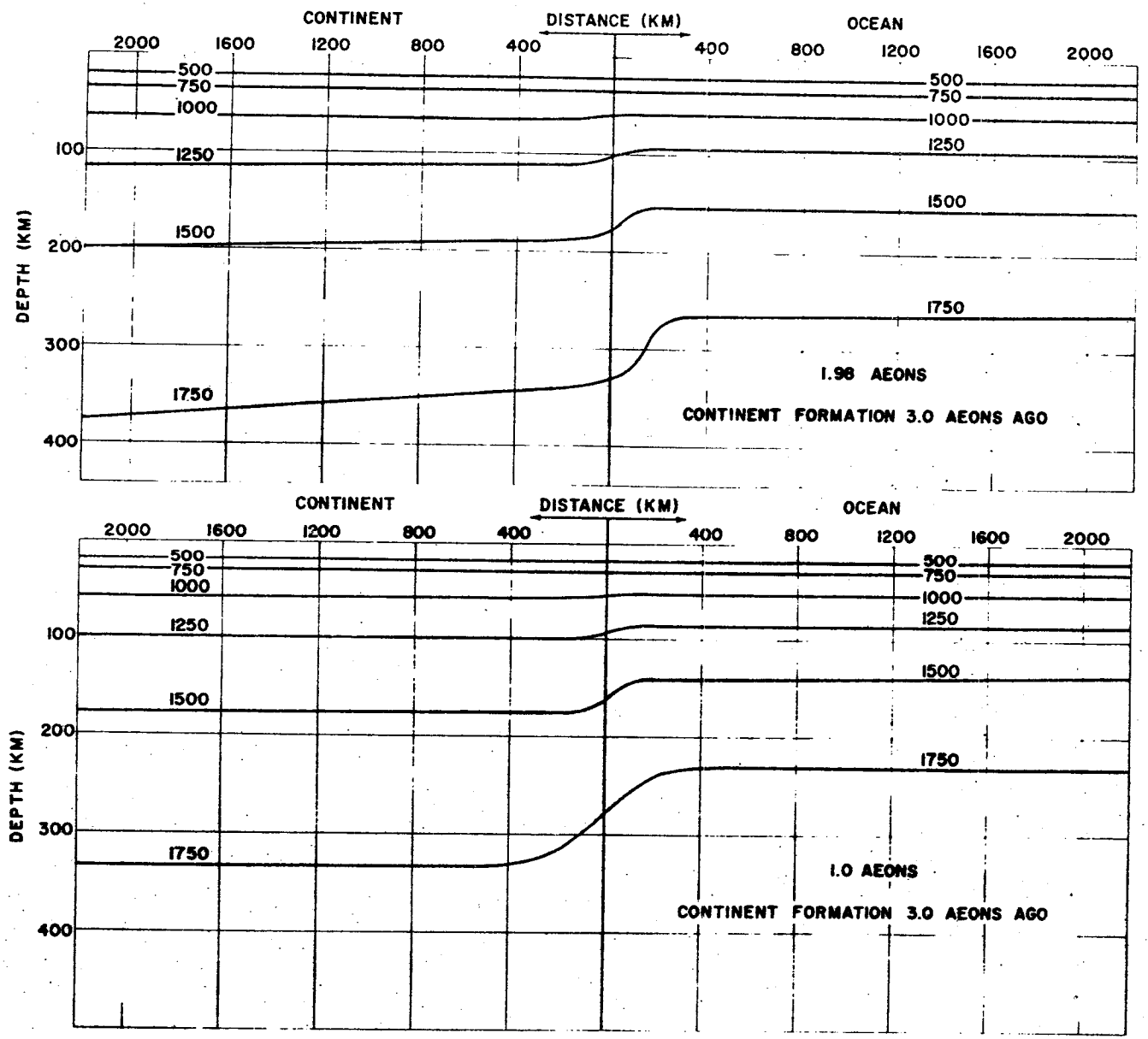


Figure 14: The development of the temperature field in a continent formed 3×10^9 years ago. Total radioactivity is that of a chondritic mantle model earth. Initial temperature is along Curve B of Figure 13.

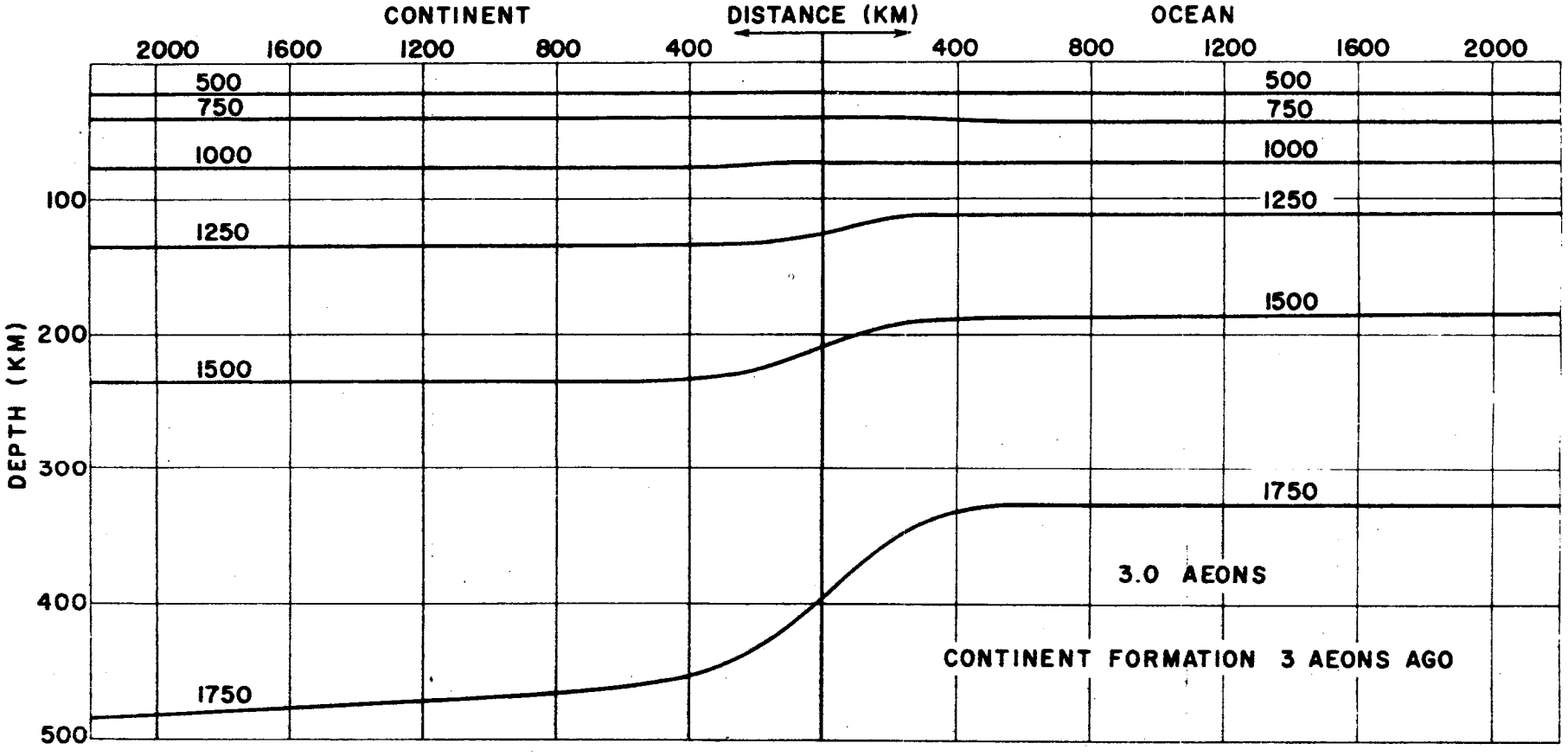


Figure 15: Further development of the temperature field illustrated in Figure 14.

Figure 16 shows the development of the heat flow for the models illustrated in Figures 14 and 15. At present, the average heat flow is about 61 erg cm^{-2} sec^{-1} and there is less than an erg difference between the heat flow under continents and under oceans. As in a number of other models that have been studied, the heat flow is high near the boundary on the continental side and low on the oceanic side as a result of the flow of heat at depth from the high-temperature oceanic regions into the lower-temperature continental regions.

Figures 17 and 18 show the development of temperature for a continent formed 3×10^9 years ago where the total radioactivity is that of a chondritic earth and the initial temperature is low (Curve B of Figure 13). The isotherms are displaced upward compared to the case where the radioactivity is that of a chondritic mantle. Under the oceans, the 1500° isotherm is at 130 km compared to 190 km for the lower-radioactivity case. The difference in depth of the isotherms between continents and oceans is slightly smaller than in the case with the lower radioactivity. The near-surface gradients are higher and this is reflected in the heat flow which is shown in Figure 19. The heat flow decreases continually with time, reaching a value of about 72 erg cm^{-2} sec^{-1} under both continents and oceans. This is about 9 erg cm^{-2} sec^{-1} greater than the observed heat flow and 11 erg cm^{-2} sec^{-1} higher than for an earth in which only the mantle has a chondritic radioactivity.

The influence of the initial temperature distribution is taken up by examining the development of temperature for the case of a mantle

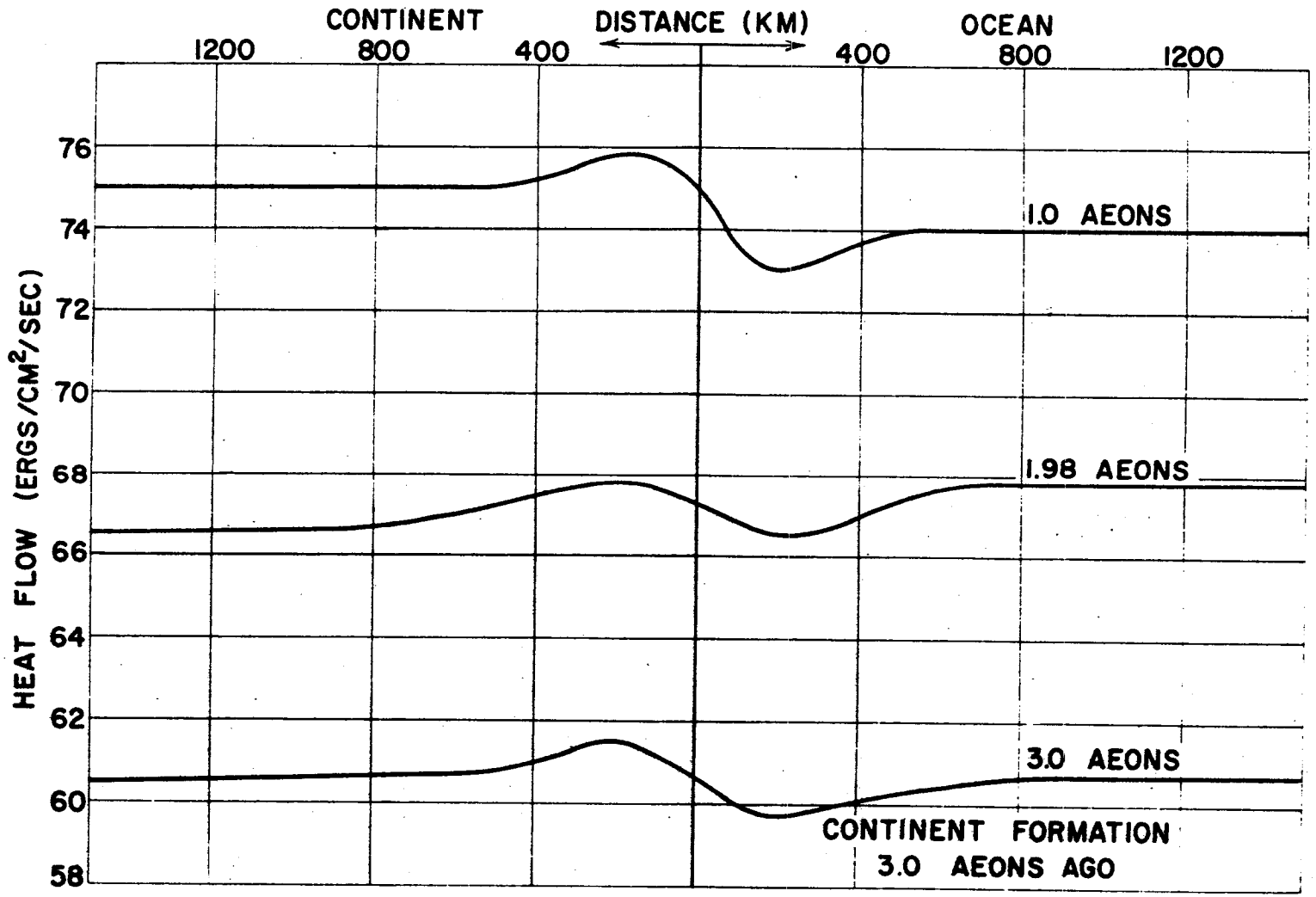


Figure 16: The development of the surface heat flow with time for the models illustrated in Figures 14 and 15.

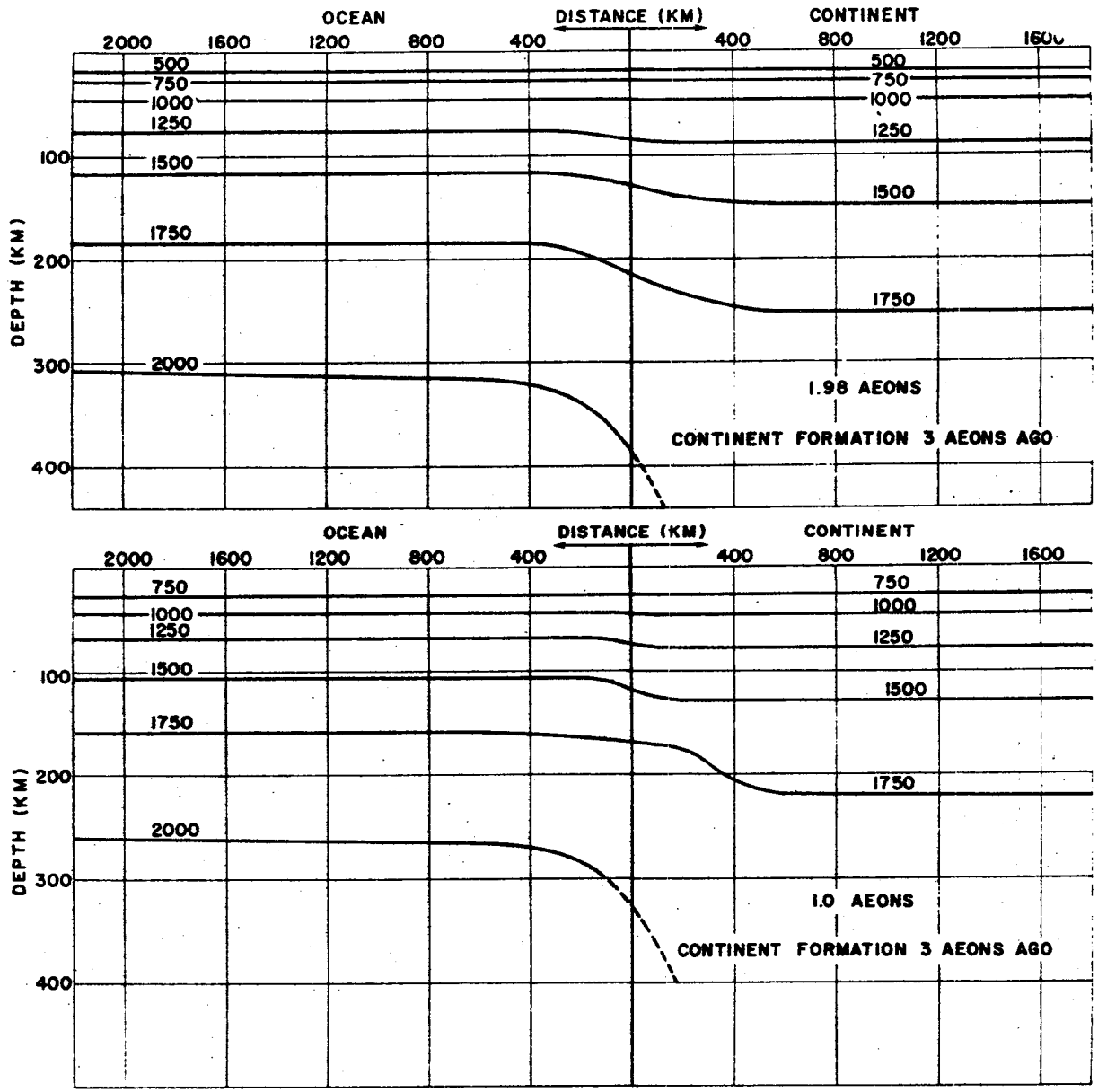


Figure 17: The development in time of the thermal field for a continent formed 3×10^9 years ago, where the total radioactivity is that of a chondritic earth and the initial temperature is along Curve B of Figure 13.

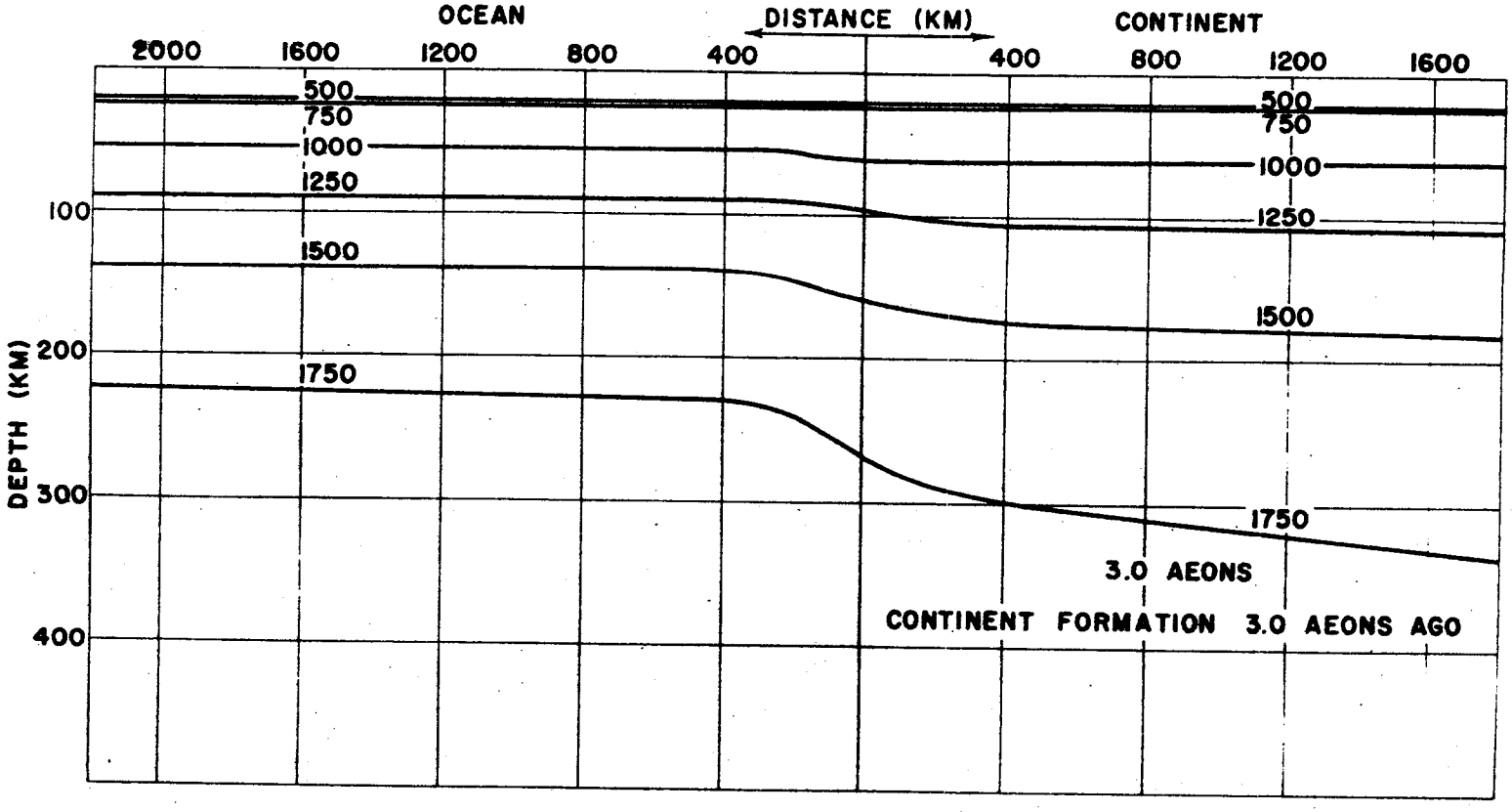


Figure 18: Further development of thermal field illustrated in Figure 17.

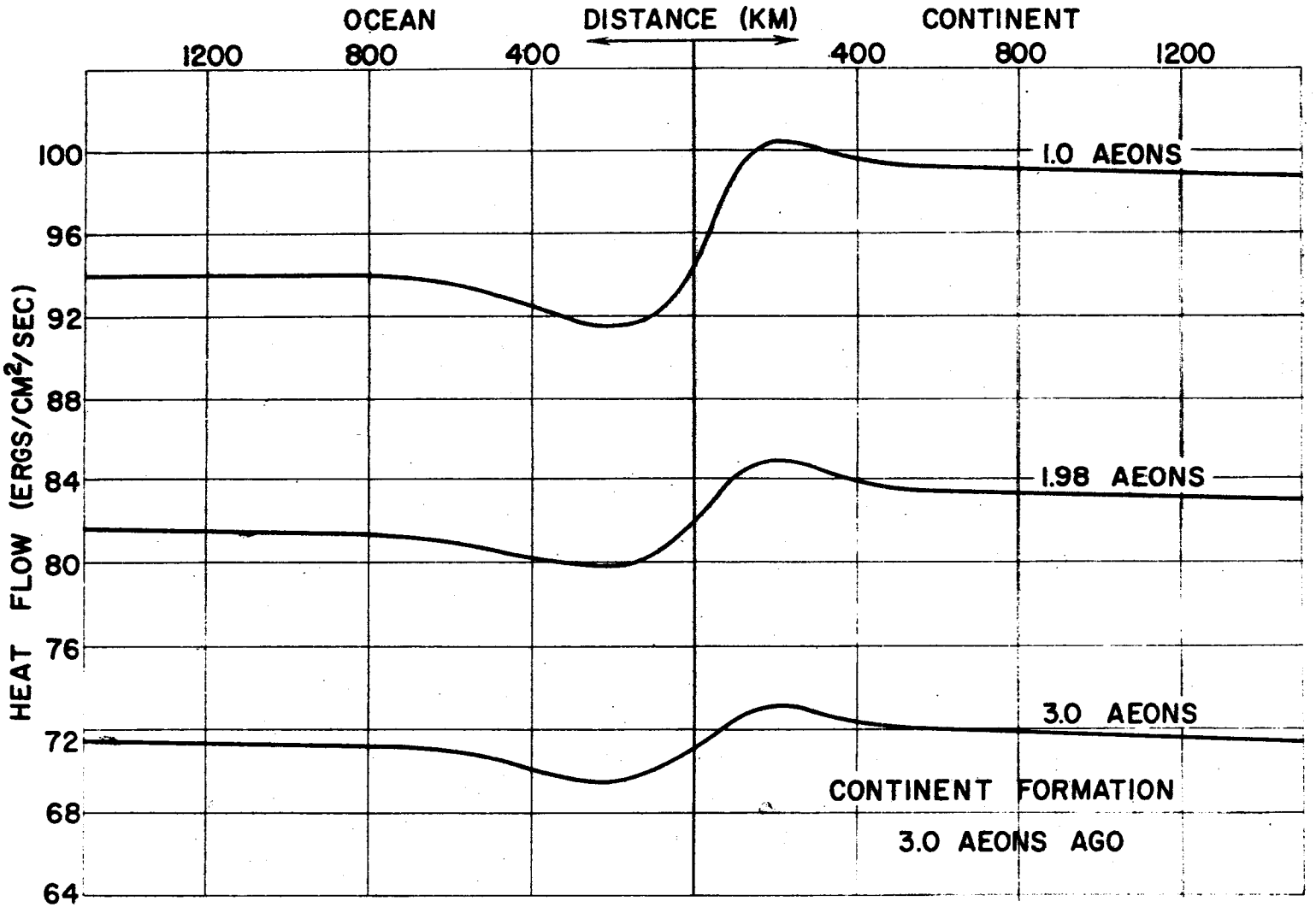


Figure 19: The development in time of the surface heat flow for the models illustrated in Figures 17 and 18.

having a chondritic radioactivity and the initial temperature on Curve A in Figure 13. Figures 20, 21, and 22 show the development of temperature as a function of time for a continent formed 3×10^9 years ago with the higher initial temperature. As in the other two cases discussed above, the isotherms progressively move down, with the 1500° isotherm moving from 110 km to a depth of 160 km below the oceans over the last 2×10^9 years. The temperature gradients between continents and oceans are somewhat less than in the case of the lower initial temperature. Figure 23 shows the variation of the surface heat flow with time. As before, there is a maximum of the heat flow on the continental side of the continent-ocean border and a minimum on the oceanic side. The mean heat flow is 64 erg cm^{-2} sec^{-1} and is nearly equal on both continental and oceanic borders after 3×10^9 years. The heat flow of 64 erg cm^{-2} sec^{-1} can be compared with the heat flow of 61 erg cm^{-2} sec^{-1} obtained with the same radioactivity, but with a lower initial temperature. The variation of initial temperature thus contributes only about 2 erg to the final heat flow.

Expansion of the Earth

It is usually assumed that the development of temperature within the earth will lead to a shrinking of the outer surface (Jeffreys, 1959). The concept of a contracting earth has played an important role in certain theories of mountain building. Undoubtedly, these theories were influenced by the demands of geologists to account for the apparent shortening of the earth's crust in mountain belts. However, the view that the earth has contracted over geologic time is based on the

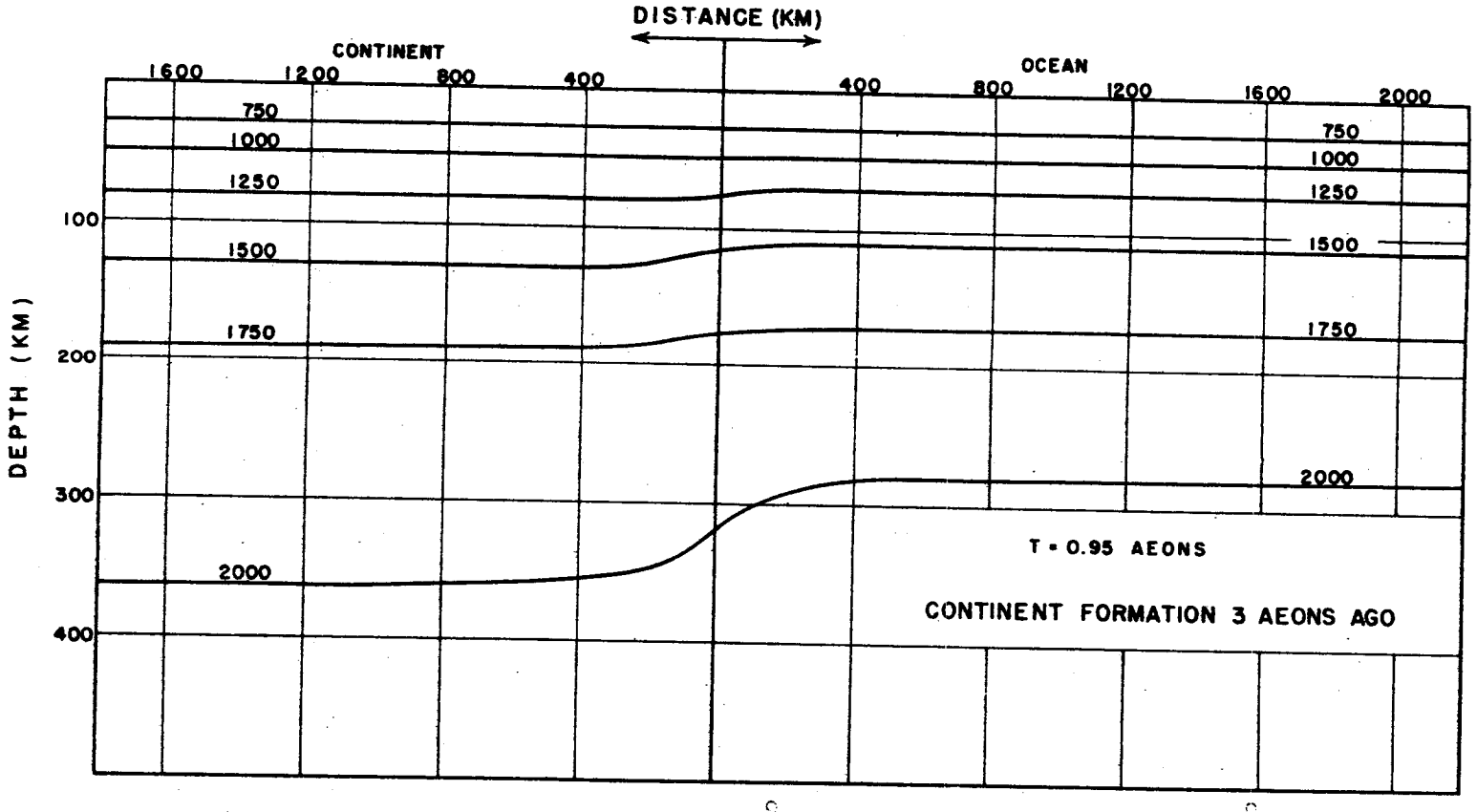


Figure 20: The temperature field after 0.95×10^9 years for a continent formed 3×10^9 years ago with initial temperature distribution along Curve A of Figure 13 and with a chondritic mantle radioactivity. The effect of the initial temperature on the development of the thermal field can be obtained by comparison with Figures 14 and 15.

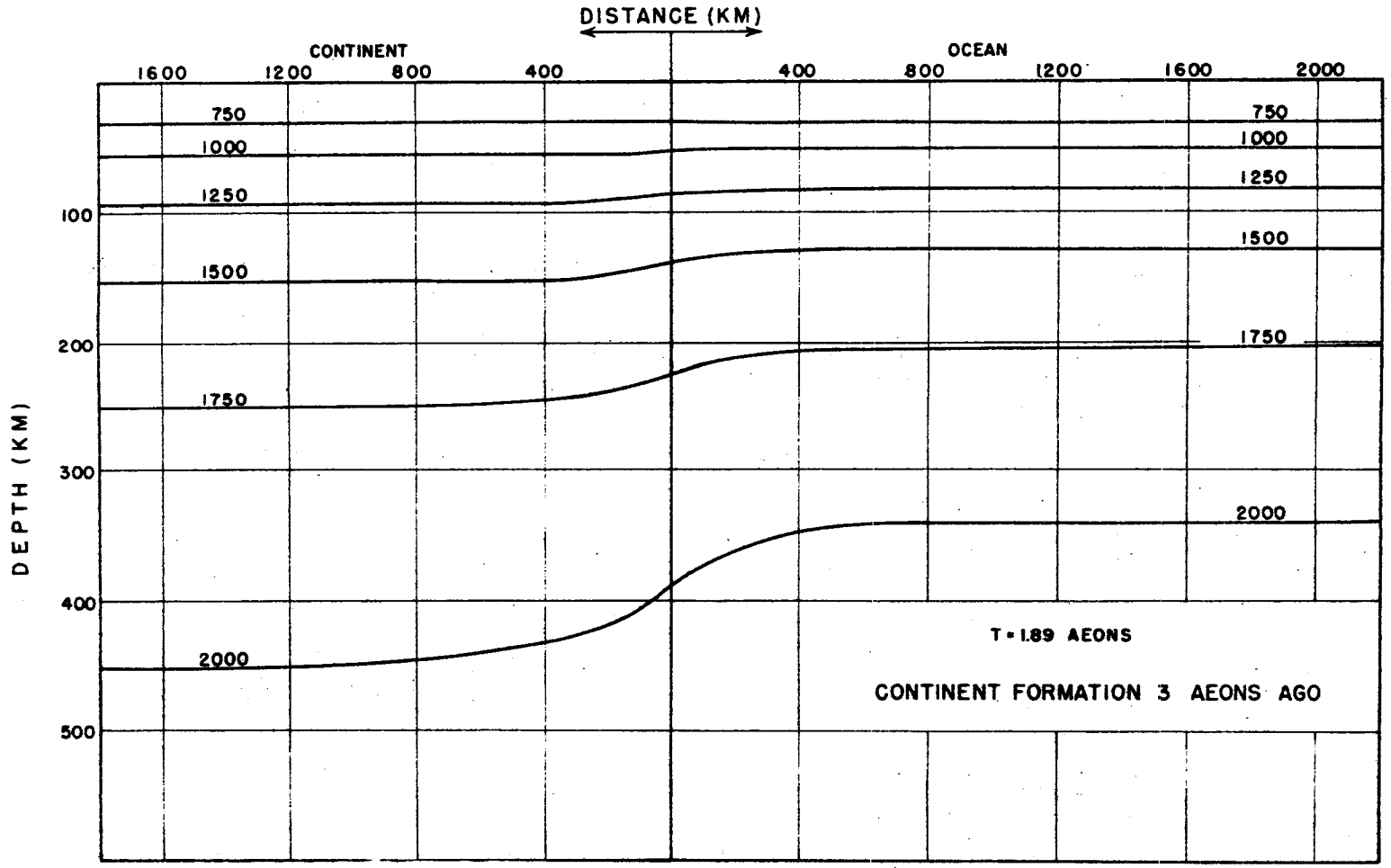


Figure 21: Temperature after 1.39×10^9 years for a continent formed 3×10^9 years ago.

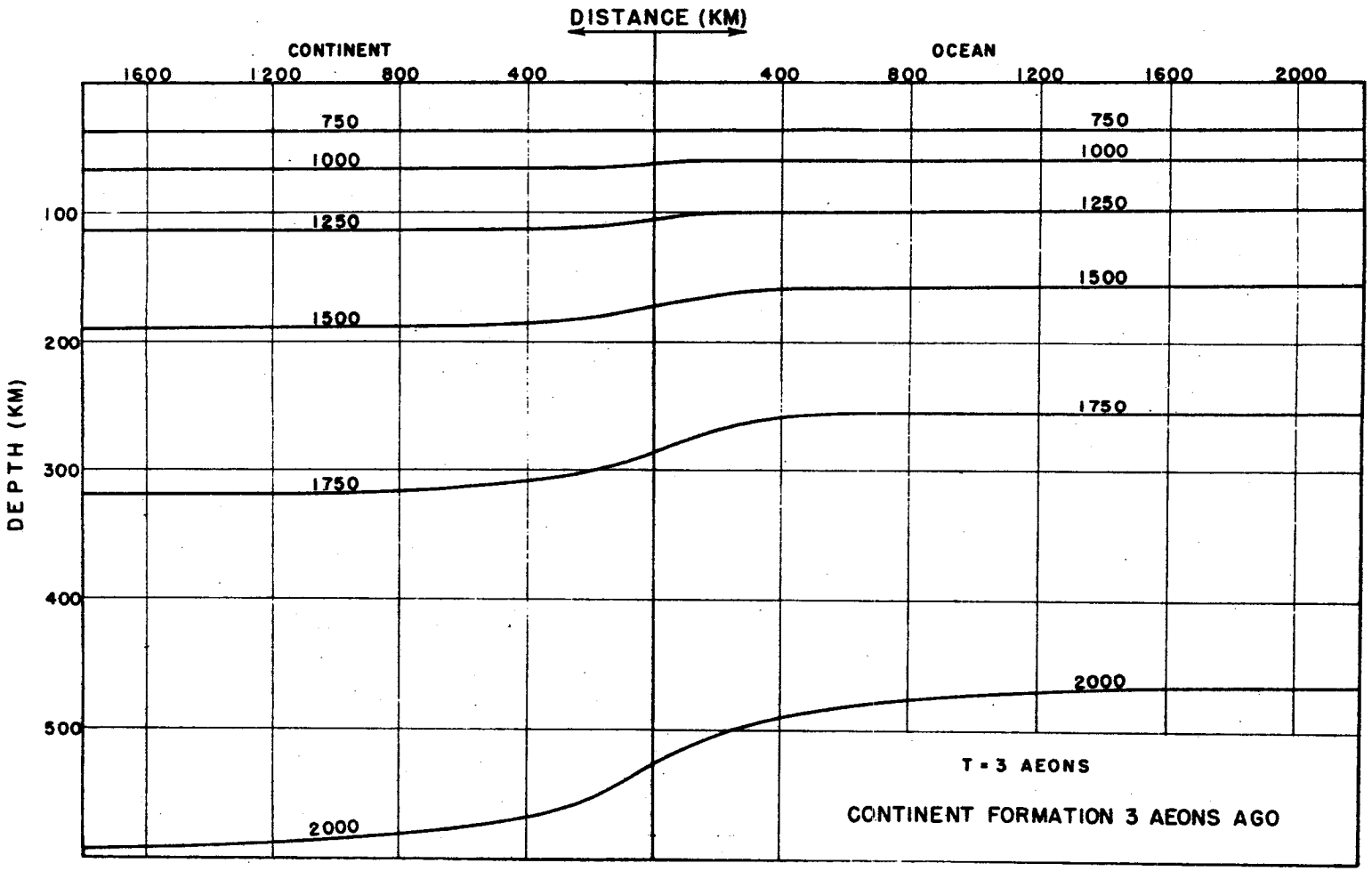


Figure 22: The present temperature distribution for a continent formed 3×10^9 years ago with an initial temperature along Curve A of Figure 13.

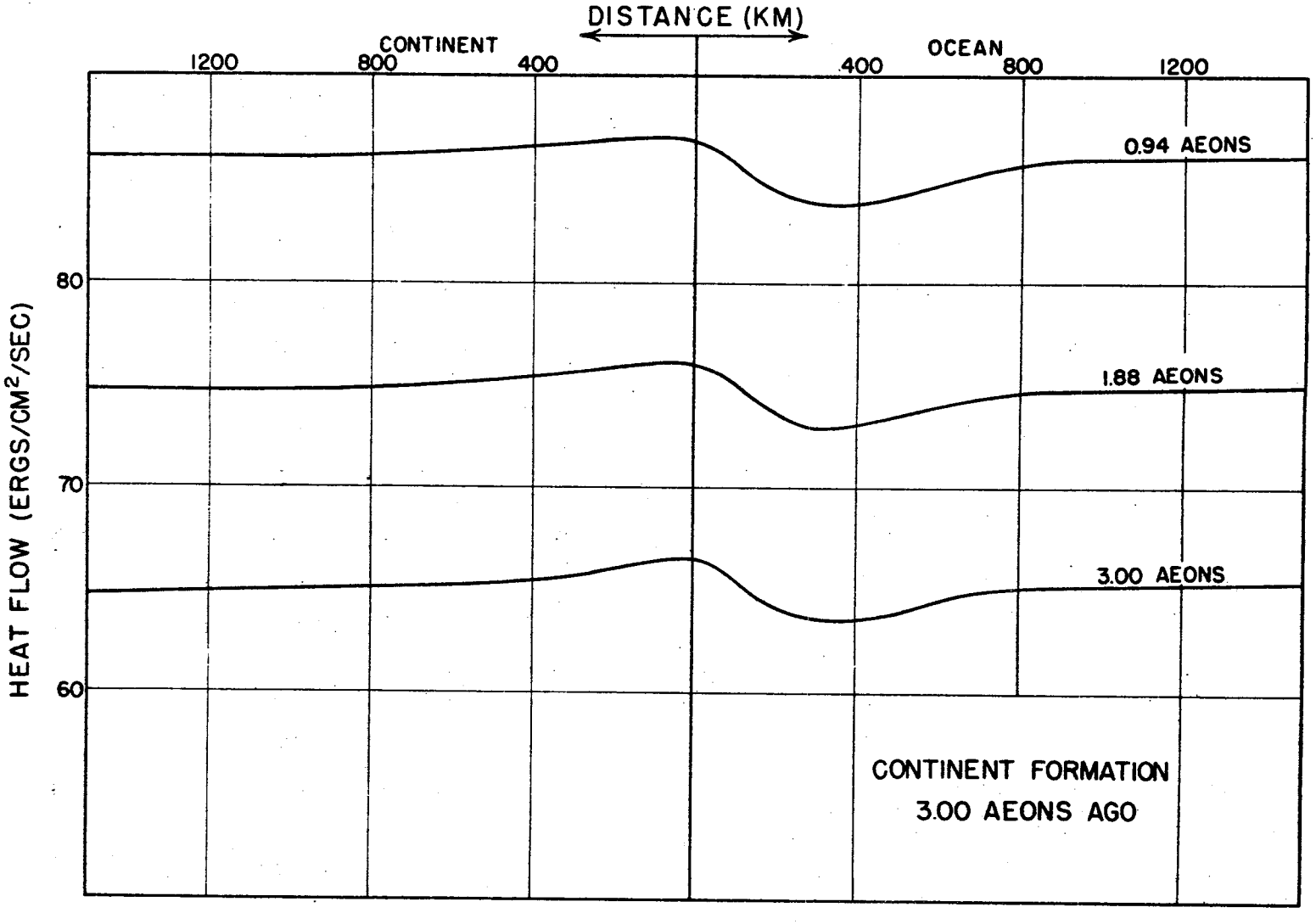


Figure 23: The development of surface heat flow for the temperature models shown in Figures 20, 21, and 22.

assumption that the earth initially was hot and has been cooling off since earliest times. Explicit account has not been taken of the effects of the variation with time on the abundances of the radioactive elements.

We suppose that the earth accumulated as a solid body with an initial temperature distribution well below the melting point of the materials. During early times, the rate of heat production was 8 times the present rate provided the earth has a radioactive composition similar to that of chondrites (see Figure 9). Because solid silicates are good thermal insulators at low temperatures, a major fraction of the heat produced would not reach the surface. The excess of the heat production over the heat loss would lead to a thermal expansion of the planet. Even if the lower portions of the mantle were raised above the initial melting temperature, the expansion would continue since only a relatively thin shell of insulating material on the outside would be sufficient to slow the rate of heat loss and the heat production would still exceed the heat flow.

It thus appears that if the earth accumulated as a solid body and if the heat-producing elements are in the ratio observed in chondrites, then the early history of the earth was a tumultuous one with the earth expanding at a rapid rate. There has been much speculation regarding an expanding earth as a consequence of a changing gravitational constant (Wilson, 1960). However, such a drastic assumption is not required since an expanding earth is a natural consequence of the thermal

development of a planet having the composition of chondritic meteorites or one in which the abundance of relatively short-lived isotopes such as K^{40} and U^{235} is high.

We consider a radial distribution of heat sources. For an elastic sphere, the outward flux of matter can be written as

$$\rho v_r \Big|_\sigma = \frac{\alpha}{c_p} \left\{ \frac{P}{R^2} - Q \right\} \tag{36}$$

α = volumetric coefficient of expansion,

$P = \int_0^R r^2 A(r) dr$,

Q = outwardly directed heat flux on the surface σ of a sphere of radius R,

(MacDonald, 1963). In the derivation of (36) it is assumed that the ratio α/c_p is constant over the volume. Both from experimental data and seismic observations it is known that this ratio does not vary by more than a factor of 2 throughout the earth.

In a sphere in which α/c_p is a constant, we see that the rate of motion of the outer surface is proportional to the difference between the heat produced per unit area and the heat flux out of the sphere. As is intuitively obvious, if the instantaneous rate of heat production exceeds the heat flux, the sphere expands due to the rise of temperature within the interior. Alternatively, if the heat flux exceeds the production rate, there is a contraction. The outward flow of heat is

a function not only of the current rate of heat production, P, but also of the initial temperature and of the past history since heat produced at a greater rate in the past may not have escaped but be contributing to the present heat flow.

In order to illustrate the numerical values involved, we consider an earth with the distribution of heat sources as given in Figures 11 and 12. Since the heat flow is very nearly the same over continental and oceanic regions, there is no need to consider the difference in the distributions between continents and oceans in discussing the thermal expansion. Taking the representative value of α/c_p of 1×10^{-12} g erg^{-1}, the rate of expansion for an earth having both a mantle chondritic radioactivity and a whole earth chondritic radioactivity is given in Figure 24. For an earth in which the total radioactivity equals that of chondrites, the initial rate of radial expansion is about 40 km in 10^9 years. After 10^9 years, this rate has been halved and after 3.6×10^9 years the earth begins to contract at a slow rate. If the total radioactivity equals that of a chondritic mantle, the rate of expansion is less. The rate of expansion is not directly proportional to the heat production. The reason for this is that the heat flow depends on the internal distribution of temperature which, in turn, is fixed in part by the thermal conductivity. The thermal conductivity is a nonlinear function of temperature. With the higher rate of heat production, the temperature in the interior is greater leading to a larger conductivity and eventually to a higher proportional heat flow.

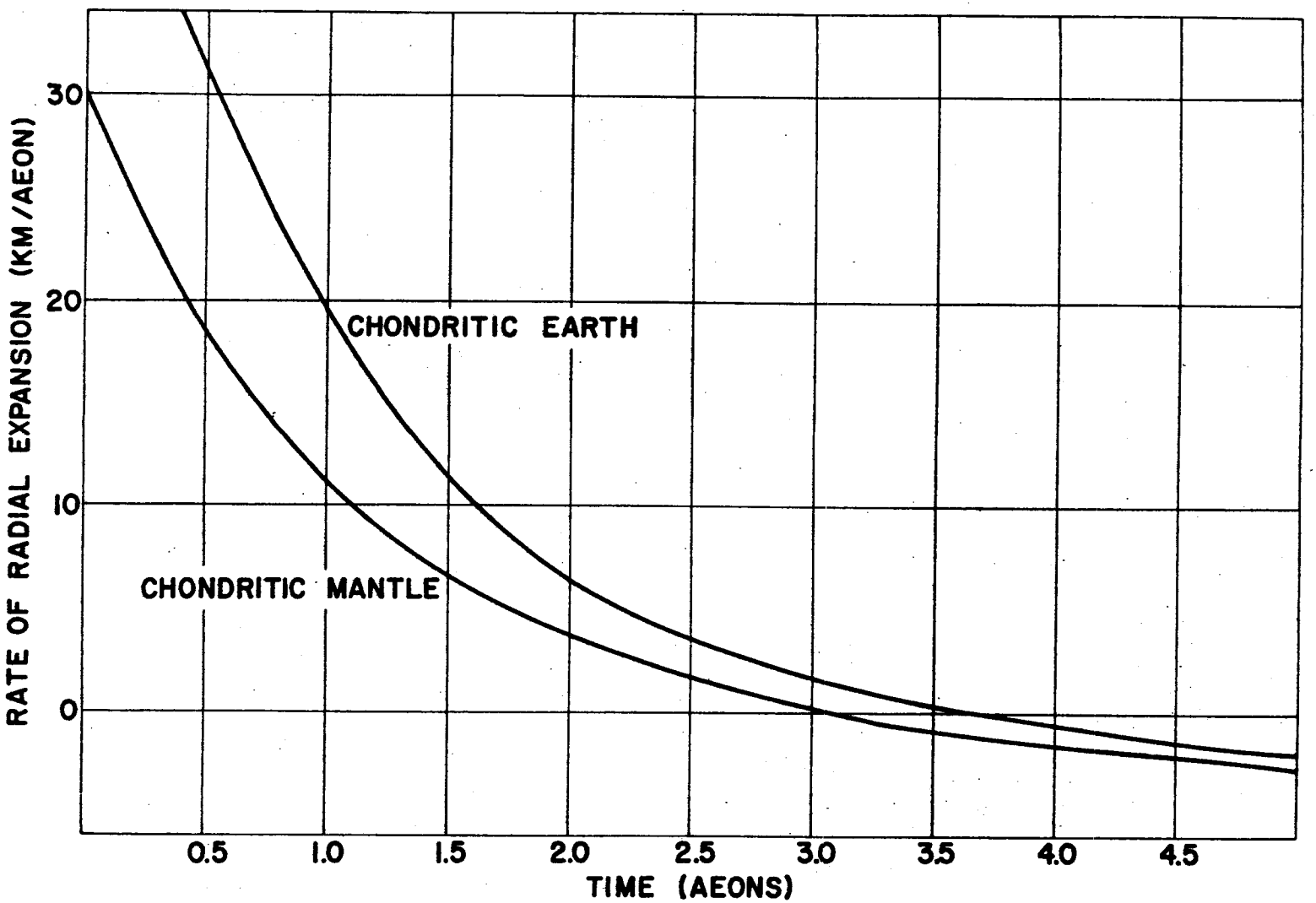


Figure 24: The rate of radial expansion in kilometers per 10^9 years for an earth having an initial temperature distribution along Curve B of Figure 13.

The curves in Figure 24 can be integrated to yield the total change of radius and this is shown in Figure 25. The present radius is taken as the reference radius. For a total radioactivity equal to that of chondritic meteorites, the radius 4.5×10^9 years was 52 km greater than it is at present.

In these calculations no account has been taken of the possibility of phase transitions within the earth. Phase transitions have the effect of amplifying the expansion since the effective thermal expansion is greater. MacDonald (1963) shows that phase transitions would have the effect of about tripling the rate of expansion.

Interpretation of the Variation of Seismic Velocities in the Mantle

Several features of the variation of seismic velocity with depth shown in Figure 2 require an explanation. First, there is the discontinuity between crustal material and material underlying the crust. This discontinuity can be interpreted either in terms of a relatively sharp change in the chemical composition of the material or in terms of a phase transition in which materials of low density transform to high-density compounds.

Below the crust, the velocity decreases with depth, reaching a minimum at a depth of 150 km. The decrease of velocity with depth can result either from the effect of changing temperature and pressure or from a change of chemical composition. Between 200 km and 900 km the rate of rise of velocity is much larger than below 900 km. The

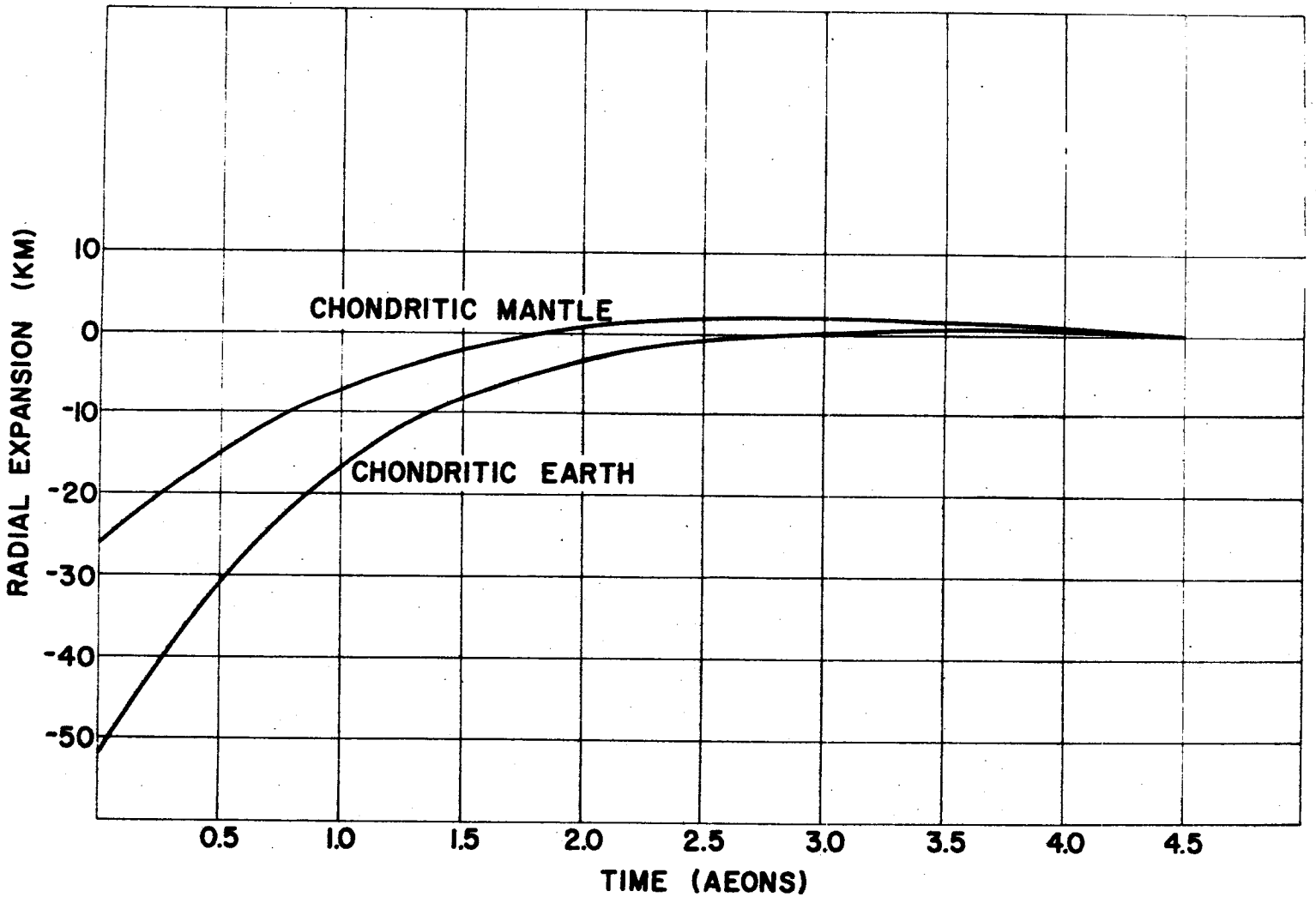


Figure 25: The radial expansion for two model earth having initial temperature along Curve B of Figure 13. The present radius is taken as the reference radius. For a chondritic model earth, the radius has expanded 51 km over 4.5×10^9 years. The calculated radial expansion is based on the constant value of α/c_p of 1×10^{-12} g erg^{-1}.

rapid rise of velocity again indicates a gradual change of composition, or of phase, or of both. The discontinuity between the elastic mantle and the fluid core can also be interpreted in terms of a phase transition or a sharp break in chemical composition. The radial variation of velocity does not take into account the near-surface inhomogeneity of material associated with the continent-ocean system. Recent investigations show that observable differences exist between the mantle under continents and under oceans (MacDonald, 1963).

Seismic investigations of crustal structures establish that the average thickness of the normal continental crust is approximately 35 km (see Table VI). The compressional and shear wave velocities are about 6.2 and 3.6 km sec^{-1} respectively near the surface and increase with depth to values about 7.0 and 4.0 km sec^{-1}. The seismic velocities for the sub-crustal mantle material center about 8.1 and 4.7 km sec^{-1}. The seismic results for the oceanic regions indicate that the discontinuity separating the crustal material from the mantle lies at depths approximately 10 to 12 km below the surface of the sea. In general, the oceanic crust is somewhat less than 5 km thick with seismic velocities for compressional waves of 6.3 to 7.0 km sec^{-1}. The rocks below the discontinuity have compressional wave velocities centering around 8.1 km sec^{-1} although a larger variation in these velocities is observed in the oceanic mantle than in the continental mantle.

The interpretation of the seismic data in terms of chemical composition cannot be unique. The seismic data must be combined with

independent data in order to obtain an interpretation of the composition of the crust and of the upper mantle. At present, the hypothesis of an oceanic crust with basaltic composition is compatible with both gravity and seismic results. Oceanic islands give no evidence for large amounts of any other material. Hence, it is generally assumed that the crust is predominently basaltic. The continental crust is inferred to be of a far more complicated character. Surface outcrops indicate a wide avariety of compositions and the observed variations of seismic velocity indicate a further variation with depth. Sub-crustal materials in both oceanic and continental regions are generally assumed to be peridotite of similar rock. In this interpretation the crust-mantle boundary is the result of a sharp change of chemical composition where the material of the crust has a greater proportion of Na, K, Ca, Al, and Si than does the sub-crustal material. In this picture, the crust developed during the early history of the earth when the earth was undergoing rapid expansion with tension fractures opening into the upper mantle. The resulting release of pressure led to melting of the lower-melting components of the mantle material and these rose to the surface through the fractured zones. The flooding of the initial surface by the basaltic material formed the initial crust. Inhomogeneities led to a concentration of material in certain areas and this, in turn, led to the formation of the early continental masses (MacDonald, 1963).

An alternative hypothesis about the nature of the seismic discontinuity between crustal and mantle material is that of a phase

transition. The phase transition would involve the breakdown of less-dense silicate phases to denser forms. Associated with the structural rearrangement there is a 10 to 15 per cent change in density. Laboratory investigations of the breakdown of low-density silicate forms are illustrated in Figure 26. The low-density feldspar albite ($NaAlSi_3O_8$) transforms to the higher-density pyroxene ($NaAlSi_2O_6$) at a pressure equivalent to depths of 400 km provided that temperature is on the order of 400°C. Other materials that are common in crustal rocks also undergo phase transitions at nearly the same conditions.

The simple suggestion that the seismic discontinuity represents a phase transition from a basalt-like material into a high-pressure form faces serious difficulties. Since basalts form a multi-component system, a transition would be anticipated to occur over a wide range of temperatures and pressures rather than along a single curve on the temperature-pressure plane. Although the width of the transition zone is uncertain, it could be on the order of a few thousand bars and if this were the case, the seismic transition zone would be spread over several kilometers of tens of kilometers. The seismic evidence on the sharpness is still unclear, but most investigators assume that the transition from crust to mantle takes place over a relatively sharp interval; a few kilometers at most.

An additional difficulty for a phase transition in oceanic areas follows from studies of the heat flow in oceanic areas. Under oceans the crust-mantle boundary is at a shallow depth, yet the 500° isotherm

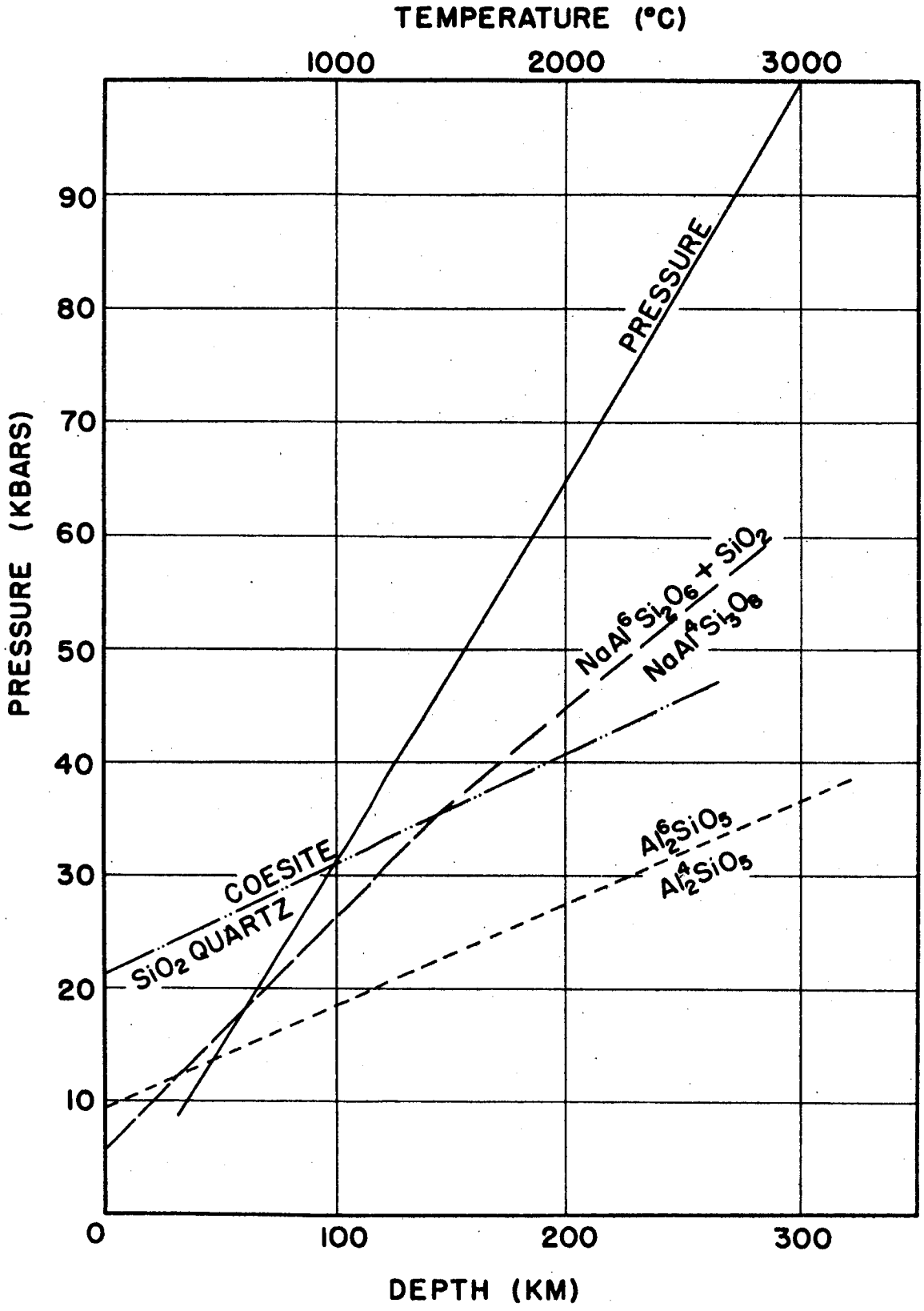


Figure 26: Stability relations between sillimanite and kyanite (Al_2SiO_5); quartz and coesite; and albite ($NaAlSi_3O_8$), jadeite ($NaAlSi_2O_6$), and quartz.

is at about the same depth under both continents and oceans (see Figures 15, 18, 22). The different depths of the seismic discontinuity can be explained in terms of a phase transition only if the crust and mantle have different compositions under oceans and continents.

Interpretation of the Low-Velocity Layer

The physical conditions which give rise to the anomalous decrease in velocity below the crust are of great interest. Laboratory measurements show that in silicates the velocity increases with an increase in pressure; pressure stiffens a rock. An increase in temperature has the opposite effect, decreasing the wave velocity. In the outer regions of the earth, both the pressure and temperature increase; velocity decreases if the increase in temperature wins out over the increase in pressure. An extrapolation of laboratory data indicates that a gradient of 6 to 7 degrees per kilometer is sufficient to produce a decrease in velocity (MacDonald and Ness, 1961). If the low-velocity zone is indeed due to the larger effect of increase of temperature, then the existence of the low-velocity zone indicates the magnitude of the temperature gradient in the upper regions of the mantle. Gradients between 30 and 150 km must be at least 6° to 7° per kilometer while below 150 km the gradients must be less. Furthermore, this interpretation would require that the seismic velocities differ under continents and oceans since the thermal gradients differ. There is a rough agreement between preliminary seismic results on continent-ocean structure and the required thermal gradients (MacDonald, 1963).

Structure of the Mantle between 200 and 900 Kilometers

The understanding of the variation of velocity between 200 and 900 km in the earth is important in any discussion of the internal makeup of the inner planets. The pressure at 200 km in the earth is 0.65×10^5 bars and this pressure is reached within Venus, Mars, and Mercury. At these depths, the variations of properties with pressure and temperature may be of greater significance than the variations due to complicated chemical processes.

The present interpretation of the seismic results is in terms of a gradual increase in velocity over the region of 200 to 900 km (see Figure 2). Earlier, Jeffreys (1936) presented a solution with the discontinuity in the velocity at a depth of 480 km. Bernal proposed that the 480 km discontinuity was due to a phase change in which silicates transform from a lower-density olivine structure to a higher-density spinel structure. Since this initial suggestion, the stability relations between olivine and spinel forms of Mg_2SiO_4 have been extensively studied by Dachille and Roy (1956, 1958, 1960) and Ringwood (1956, 1958a,b) by examining the temperature and pressure stability of solid solutions of $Mg_2(Si,Ge)O_4$ and $Fe_2(Si,Ge)O_4$. Dachille and Roy (1960) find that the maximum silicate content of the spinel solid solution increases steadily with pressure from about 10 mole per cent at 700 bars to 50 mole per cent at 60,000 bars at 540°C. They extrapolate these data to obtain the transition of the pure end member Mg_2SiO_4 at 100,000 bars with an uncertainty of at least 15,000 bars. Ringwood (1958c) synthesized the spinel form of Fe_2SiO_4 at 38,000 bars and 600°C.

The transition of $MgSiO_3$ from a pyroxene to a corundum structure has been investigated in a similar way by Ringwood and Seabrook (1962). From these results, it is clear that at depths on the order of 300 km and greater, the common olivine and pyroxenes are no longer stable and will give way to denser phases.

A further possibility involves the conversion of olivine or spinel forms of magnesium iron silicates to oxide phases. Birch (1952) considered this possibility while noting that the ratio of incompressibility to density for a close-packed metal oxide such as MgO approached the observed ratio within the earth at depths on the order of several hundred kilometers. MacDonald (1962) has calculated the stability relations between Mg_2SiO_4 and MgO and SiO_2. At high pressures, SiO_2 inverts to a much denser form, having a density of 4.35 compared to the density of quartz of 2.7. The estimated phase transitions for Mg_2SiO_4 are shown in Figure 27.

The temperature-pressure conditions under which the high-pressure transitions take place are still most uncertain. However, it is clear that under conditions existing at depths on the order of 200 - 500 km within the earth, the common silicate phases undergo transitions to new phases with structures in which the oxygen atoms lie in closest packing. Since the oxygen atoms occupy almost the entire volume within the silicate, it is unlikely that the silicates will undergo further structural rearrangements. Any further decrease in volume can only result from a distortion of the outer electron orbits, i.e., pressure

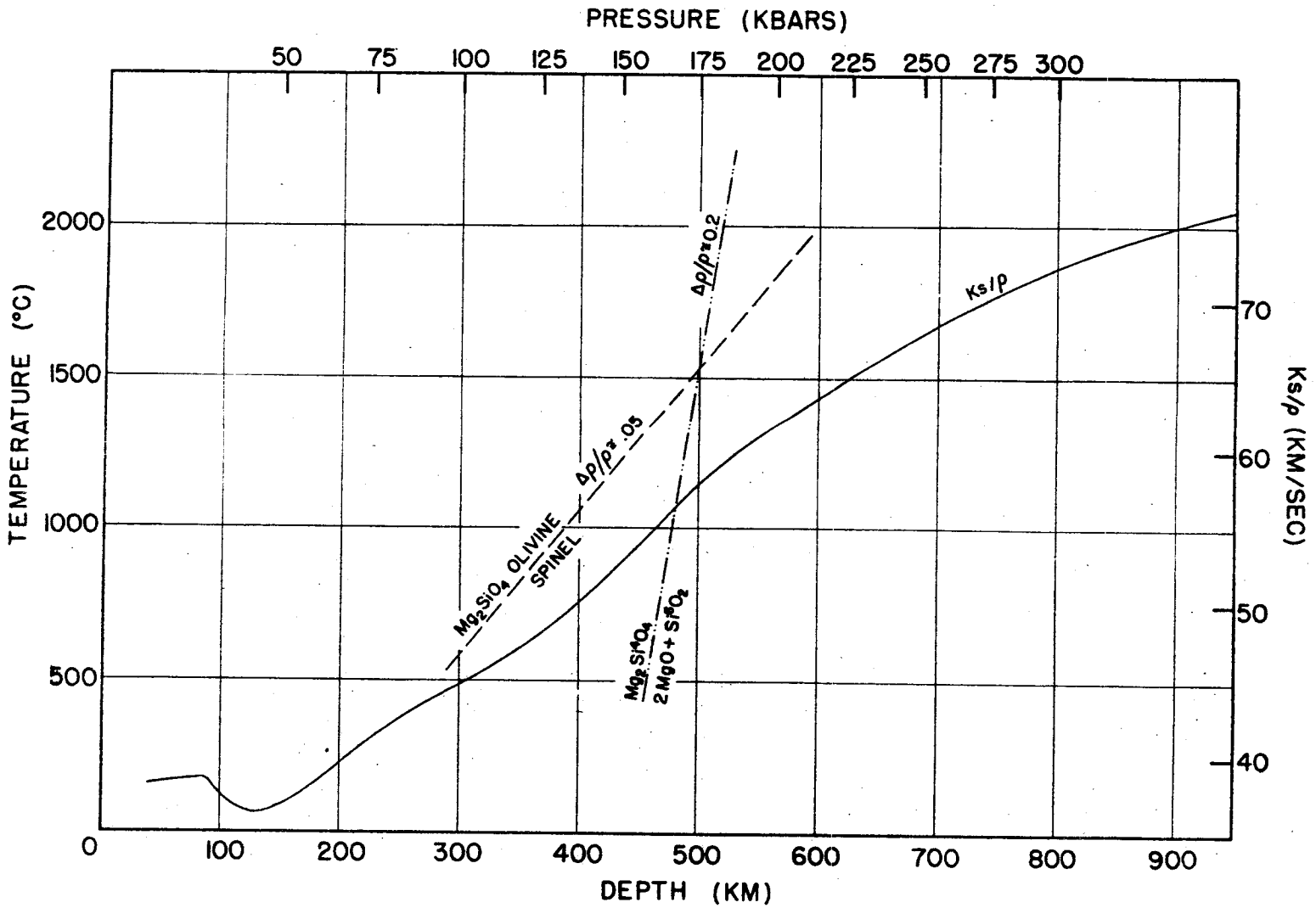


Figure 27: The calculated curve for the breakdown of olivine to periclase plus stishovite. The variation of $\varphi = K_s/\rho$ is derived from the Gutenberg model earth illustrated in Figure 2.

ionization. The pressure required for such transitions is on the order of 10^6 bars or greater and is reached only within the cores of the larger planets. The density changes in these transitions are likely to be small compared with density shifts in the geometric rearrangements. The collapse of silicate structures at a pressure of about 10^5 bars, however, can take place in Mars, Venus, and Mercury since pressures of this magnitude are reached within the planets.

Chemical Composition of the Earth's Core

The discovery of the core of the earth by Oldham in 1906 has since provoked much speculation as to its physical and chemical properties. The difficulty in determining with any degree of definiteness the nature of the earth's core is connected with the difficulty of compressing materials to the pressures realized within the earth. The pressure at the core-mantle boundary is 1.3×10^6 bars. This pressure is much in excess of the highest pressures that can be obtained in static experiments at the surface. It is possible to reach these pressures for short periods of time during the shockwave experiments and data on the compressibility of solids obtained using shockwaves are of great importance in obtaining the equation of state for the core.

The core of the earth is generally assumed to be a liquid iron-nickel alloy. The evidence usually advanced in support of this view includes:

1) the existence of iron meteorites,

2) the fact that the density of iron is in general agreement with the density of the core as estimated from seismology and from the theory of the figure of the earth, and

3) iron would probably be liquid at the temperatures and pressures of the core while silicate materials would be solid.

In addition, recent investigations of the abundances of elements in the solar system suggest that iron is an abundant heavy element and a planet with an iron core would have the proper heavy-element abundances.

The hypothesis of an iron core has not gone unchallenged. Kuhn and Rittman (1941) propose a core made up of undifferentiated solar matter, arguing that an iron core is inconsistent with the relative abundances of the elements. Ramsey (1948, 1949, 1950) suggests that the seismic discontinuity at a depth of 2900 km represents a phase change from $(Mg,Fe)_2SiO_4$ in the form of an ionic compound to a metallic phase. Birch (1952) critically reviews these suggestions and concludes that the core is principally iron-nickel with minor amounts of silicon and carbon since the apparent density of the core is somewhat less than the expected density of pure iron. The view that the 2900-km discontinuity is a phase change has many attractive features. However, a phase change in a multi-component system will almost certainly be spread out over a range of pressures. The sharpness of the 2900-km discontinuity plus the large change in density of the core-mantle boundary favors a compositional discontinuity rather than a change of phase.

The mean density of the core and the density of the core at the core-mantle boundary follows from the incompressibility obtained from seismic velocities and from the mass and moment of inertia. The satellite determinations of the earth's growth field yield a much improved estimate of the mean moment of inertia and reduce the moment of inertia by about 0.9 per cent from that previously accepted. The new value of the moment of inertia implies a somewhat greater central concentration of mass. Using the velocities given in Figure 2 and the gravitational data listed in Table I, the mean density of the core is 11.3 and the density at the core-mantle boundary is 10.2.

Shockwave measurements on the equation of state of iron (Knopoff and MacDonald, 1960; Birch, 1961) indicate that the density of iron at 1.3×10^6 bars at about 2000°C is 11.2 g cm^{-3}. This value should be compared with the core boundary density of 10.2. The core material is thus about 10 per cent less dense than iron. Nickel would raise the mean density, while the lighter elements, in particular silicon and carbon, would lower the density. MacDonald and Knopoff (1958) suggest that silicon is the principal alloying element in the core.

The Earth's Magnetic Field

The magnetic field of the earth resembles to a first approximation that which would be produced by a dipole at the center. The dipole is inclined at an angle of 11.4 degrees to the axis of rotation. The vertical field at the equator is 0.31 gauss and 0.63 at the geomagnetic pole. The field at the earth's surface can be shown by a spherical harmonic analysis to consist of two parts:

1) that due to sources located within the earth,

2) that having an origin outside the earth.

In the following, we will be concerned with the dominant part of the field which has its origin at depth.

A detailed description of the earth's field shows that it differs from that of a dipole in a substantial fashion. There are also large-scale variations of the field whose origin must lie at depth.

The terms in a spherical harmonic analysis of the field can be interpreted as hypothetical multipoles located at the earth's center. The lowest order multipole is a dipole which has a component along the earth's rotation axis and two mutually perpendicular dipoles lying in the equatorial plane. Beyond the dipole are the five quadrapoles. Table XIV lists the dipole and quadrapole component obtained by Vestine, et al., (1947). The striking feature is the predominance of the dipole component along the rotational axis as compared to the components in the plane perpendicular to the axis. In addition, the quadrapole components are significant. The estimates of the higher-order moments are more uncertain, but it appears that they decrease in magnitude.

The magnetic field shows both high-frequency and low-frequency variability in time. If the short-period fluctuations are filtered out, there remains a slow change called the secular variation. This change usually continues at a point for many years and the rate of

TABLE XIV

Multipole Components of the Earth's Magnetic Field
(Vestine et. al, 1947)

(Units of 10^{-4} gauss)

axial dipole	equatorial dipole	quadrapole components
-3057	-211	-127
	581	+296
		-166
		+164
		+ 54

change of any component of the field is usually less than one per cent per year. The changes involve all components of the field. The main features of the non-dipole field drift westward at a rate of about 0.2 degrees per year. The dipole field itself moves more slowly but it is not known whether the motion is systematic.

A large number of studies have been carried out on the field in ancient times. The critical question of the variation of the strength of the field with time has not been resolved. Gauss carried out an analysis in 1835 in which he found that the axial dipole had a component of 0.323 gauss which can be compared with the 0.306 gauss determined by Vestine. This decrease in the field strength appears real. Further there is evidence that in Roman times the strength of the field was 1.5 its present value. The average rate of decrease in the dipole field since Roman times has then been much less than that based on direct observation. This would suggest that there is much variability in the field even over time-scales measured in hundreds or thousands of years.

There is statistical evidence that the time average of the geomagnetic field was an axial dipole during much of the Tertiary (Torreson, Murphy, and Graham, 1949). It is now generally accepted that at certain periods the dipole may have been reversed. This is based on many rock samples from different localities that have shown a magnetic polarization opposite in direction to that of the present field. It is not possible to say how many reversals of the dipole

have taken place, but the time-scale between reversals may be short compared with a million years.

While the evidence in recent geologic times is for a dipole field, there is a great deal of data on remnant magnetization that suggest that the field either was not a dipolar field in the past or that the field has changed directions relative to the axis of rotation. The complicated questions regarding the analysis and interpretation of the ancient geomagnetic field are taken up by Cox and Doell (1960) and Runcorn (1962a).

The Origin of the Magnetic Field

The origin of the geomagnetic field is generally thought to lie in motions within the core. The short time-scale excludes a possible origin in the mantle since there is no evidence for short time-scales for processes within this part of the earth. The lack of correlation of the field with geologic features rules out an origin in the crust. The core, on the other hand, provides a suitable place for relatively rapid motions to take place. The core is known to be a fluid and, therefore, rapid relative motions of the material relative to the mantle and crust can take place. In addition, it is probably an electrical conductor since both chemical arguments and the density imply an iron-nickel alloy. Given an initial seed field, the motions induced in some fashion within the conducting fluid could give rise to a self-sustaining dynamo. The conducting fluid would carry the lines of current flow and the complexity of the terrestrial field

could be explained by the complexity of the fluid motion. A further general feature of the magnetohydrodynamic origin of the field is the role of rotation. The only force comparable in size to the Lorentz force is the Coriolos force. The rotation influences the geometry of the resulting field, but does not enter into the energetic considerations.

While it is generally agreed that the field originates in the fluid motions of the core, the cause of the motions is a matter of dispute. Elsasser (1939) and Bullard (1949) conclude that the only likely cause is thermal convection due to radioactive heating of the core. In order to maintain a field, the radioactivity of the material of the core need only be a few per cent of the radioactivity found in surface rocks, but be substantially greater than that found in ultrabasic rocks, and several orders of magnitude greater than the radioactivity observed in iron meteorites. In order to overcome this difficulty, there have been various suggestions (Verhoogen, 1961) that the heat source lies within the solid inner core and may be due to the condensation of the fluid phase into a solid form.

An alternative approach to the theory of the earth's field has been put forward by Malkus (1963). Malkus reviewed the theory that the motion is produced by the difference in precessional torques acting on the core and mantle of the earth. The precessional torque is proportional to $\frac{C - A}{C}$, where C and A are the principal moments of inertia. Because of the differing distribution of density, the

dynamical flattening of the mantle and core differ and, as a result, the precessional torques acting on the two parts of the earth also differ. The differential torques can then transfer momentum into the fluid motion of the core. In essence, the theory is an astronomical one. In the theory, a planet must be rotating and there must be an external body producing the torque. In the earth the energy for the motion comes not out of the radioactivity of the core, but out of the mechanical motion of the earth-moon-sun system. The theory has not been developed in detail, but Malkus shows that relatively short time constants are possible.

III. INTERNAL CONSTITUTION OF THE MOON

Almost all inferences regarding the internal constitution of the moon follow from data on the orbital and rotational motion of that body. A possible additional limitation to the thermal constitution is derived from a comparison of the emission temperature of the moon at radial and infrared wavelengths.

The mean density of the moon is well-established at 3.34 g cm^{-3}. This density is about the same as the density of upper mantle material and lower than the density of chondritic meteorites. The actual density should be corrected for the effects of temperature and pressure when such comparisons are made. Urey (1960) estimates that the density is in the neighborhood of 3.4; significantly less than the density of chondrites (3.57 - 3.76). He has pointed out that the moon's density is consistent with a material having about half the total iron content of chondritic meteorites and has considered several chemical models. It should be noted that the maximum pressure reached within the moon is about 46,000 bars. The pressure is too low for either the olivine-spinel or olivine-oxide transition to take place. However, the transitions in which the silicates undergo minor structural transitions are possible.

The Moon's Gravitational Field

Urey (1952) first emphasized that the figure of the moon sets limits to the possible internal constitution. The basic argument is that the present gravitational figure of the moon is not in hydrostatic

equilibrium. The inequalities in figure must be supported either statically by internal strength or dynamically by convective currents. If the moon were partially molten or if a large part of it were nearly melting, then the figure should be close to hydrostatic. The possible inferences of the internal thermal state from the figure are complicated by the fact that the stress differences may be supported in a thick, cold outer shell while the inner part of the moon may be molten or nearly so. However, it is difficult to provide a distribution of radioactivity that would lead to such a thermal configuration in a body the size of the moon. This point will be discussed in detail later.

It is usual to describe the gravitational field of the moon in terms of ratios of the principal moments of inertia since these ratios are determined by measurements of the moon's libration. We denote the greatest moment of inertia about the axis of rotation by C; the least moment of inertia, A, is about the axis pointing toward the earth; while the intermediate moment, B, is about an axis at right angles to the radius vector and axis of rotation. The best determined ratio is $\beta = \frac{C - A}{B}$, since it depends on the inclination of the axis of rotation to the pole of the orbit. Jeffreys (1961) and Habibullin (1958) obtain $\beta = 0.000\ 627\ 9 \pm 0.000\ 001\ 5$ (see also, Kaula, 1963e).

The ratio $\gamma = \frac{B - A}{C}$ can be derived from observations of the forced physical libration of longitude which arises from the attraction of the earth for the moon's equatorial bulge. Jeffreys (1961), using Yakovkin's (1952) data, obtains $\gamma = 0.000\ 204\ 9 \pm 0.000\ 000\ 9$. There is an ambiguity

due to the difficulty in determining on which side of the resonant point the largest term in the libration falls; the above value for γ depends on assuming that $\frac{\beta - \gamma}{\beta} = 0.67$.

In principle, it is possible to deduce the quantities $\frac{C - A}{M_{\mathbb{D}} R_{\mathbb{D}}^2}$ and $\frac{B - A}{M_{\mathbb{D}} R_{\mathbb{D}}^2}$ from observations of the secular mode of the node and perigee of the moon's orbit. In order to carry out such a determination, it is necessary to evaluate the solar contribution to the secular motion of perigee. The solar contribution is about $1.46'' \times 10^5$ per year while only 1" per year is due to the moon's figure. Therefore, a small error in the evaluation of the solar contribution can result in a large uncertainty in the flattening of the moon and the ellipticity at the equator. Jeffreys (1961) obtains

$$\frac{C - A}{M_{\mathbb{D}} R_{\mathbb{D}}^2} = 0.000\,364 \pm 0.000\,075 \quad ,$$

$$\frac{B - A}{M_{\mathbb{D}} R_{\mathbb{D}}^2} = 0.000\,071 \pm 0.000\,039 \quad ,$$

$$\frac{C - \frac{1}{2}(A + B)}{M_{\mathbb{D}} R_{\mathbb{D}}^2} = 0.000\,328 \quad .$$

These can be combined to yield

$$\frac{B}{M_{\mathbb{D}} R_{\mathbb{D}}^2} = 0.58 \quad , \tag{37}$$

$$\frac{C}{M_{\mathbb{D}} R_{\mathbb{D}}^2} = 0.35 \quad ,$$

where the ratio $\frac{B}{M_{\mathbb{D}}R_{\mathbb{D}}^2}$ is the better determined of the two values. For a homogeneous sphere, the two ratios in (37) should equal 0.4. If the results are accepted at face value, then the moon is more dense in the outer part than in the interior. Alternatively, there are errors in properly subtracting the solar contribution to the secular motion of perigee.

There are also difficulties in the geometric shape of the moon. The mean radius of the visible circle of the moon's face is 1737.85 $\pm$ 0.07 km. Alexandrov (1960) assumed the physical surface to be an equal potential and then used the mean radius with the mechanical ellipticities to deduce that the 3 radii are

$$\begin{aligned} a &= 1738.57 \text{ km}, \\ b &= 1738.21 \text{ km}, \\ c &= 1737.49 \text{ km}. \end{aligned} \tag{38}$$

Observers generally agree that the center of the visible face of the moon as derived from limb observations is about 1 km south of the center of mass of the moon as defined by the motion of the moon in its orbit. Recently, O'Keefe and Cameron (1962) suggested that the geometrical center of the moon's surface is actually 2.4 km south of the gravitational center.

The observed values of β and γ differ from those corresponding to hydrostatic equilibrium. For hydrostatic equilibrium, β_{Th} = 0.000 037 5

and $\gamma_{Th} = 0.000\ 028\ 1$. Thus, the observed β is almost 17 times as large as the theoretical β and the observed γ is 42 times as large as the theoretical γ. This has suggested to Urey (1952, 1962) and MacDonald (1961, 1962) that the moon is an elastic body capable of supporting stress differences on the order of 20 bars. However, Kopal (1962), arguing on the basis of a presumed analogy with the earth, suggests that the inequalities in figure are the result of convection currents in the interior. This point has been further considered by Runcorn (1963).

Jeffreys (1959) proposes that the departure of the figure from equilibrium is due to the solidification of the moon at a time when it was much nearer to the earth. The present figure then represents a frozen tidal wave. However, on the tidal wave hypothesis, the ratio of γ/β is independent of the distance of the moon from the earth and has a theoretical value of 0.25. The observed ratio is 0.64. Thus, the present figure cannot be explained solely by supposing that the moon solidified when it was much closer to the earth. There are further difficulties posed by the thermal history of the moon. If the moon were initially molten, a uniformly distributed radioactivity only one-fourth that of chondrites would keep the moon near or at the melting point. In this case, the frozen tidal wave could not be maintained.

Because of the uncertainties in the gravitational data, there is no answer to the question of whether or not the moon is differentiated.

An experiment of first-order importance is the placing of a satellite in orbit about the moon to determine J_2. The indications are that the moon supports stress differences of the same order as the stress differences supported within the earth. This is consistent with a finite creep strength for the material making up the moon.

Tides on the Moon

Both the earth and the sun produce significant tides on the moon. If the moon were spherical, the effect of these tides would be to distort the sphere into an ellipsoid whose long axis is directed toward the tide-raising body. The moon rotates once per month relative to the line joining the centers of the sun and moon and, therefore, the long axis of the solar tidal ellipsoid rotates once per month relative to the axis fixed within the moon, giving rise to a predominantly fortnightly tide similar to the semidiurnal tides on the earth. The moon, however, does not rotate relative to the line joining the centers of the earth and moon, but wobbles about this line with a range of about 7 degrees. The tides due to the earth are, therefore, due to two causes: the changing distance of the earth which causes the ellipticity of the tidal ellipsoid to increase as the moon comes closer to the earth and to decrease as the earth recedes; and the moon's wobble about the line of centers. The relative importance of these effects varies with position on the moon's surface. Close to points of intersection of the line of centers with the moon's surface the variation of distance is the more important. The monthly variation of the earth-moon distance is the most important factor leading to the large monthly tides.

The response of the moon to the tide-raising potential can be determined by evaluating the Love numbers for the moon. In order to determine the Love numbers, we require estimates of the density and rigidity of the moon's interior. Since the moon's density is close to that of the upper part of the earth's mantle, this suggests that the elastic properties should be similar to those of mantle material. For a homogeneous, incompressible, spherical body the Love numbers are determined by

$$h = \frac{5f}{2f + 1} \quad , \quad k = \frac{3f}{2f + 1} \quad ; \quad f = g \rho \frac{R}{19\mu} \quad , \tag{39}$$

where

g = surface gravity,

R = outer radius,

μ = rigidity,

(Love, 1927). If the density is taken as 3.34 and a rigidity of 7.4×10^{11} dyne cm^{-2} is used, we obtain $h = 0.0199$, and the tidal gravity variation $(1 + h - \frac{3}{2} k)$ is increased by 3.3 parts in 1000 as compared with the variation on a perfectly rigid moon. If the moon has an inner fluid core with a core radius equal to one-half the outer radius then $h = 0.0639$, $k = 0.0384$, and there is a change in the gravitational yielding of the moon by 6 parts in 1000.

Because of the small dimensions of the moon, the additional deformation due to elastic yielding is small compared with that of the earth. As a result, the investigation of the mean elasticity of the moon using gravimetric techniques is marginal.

Seismic Activity on the Moon

The seismic exploration of the moon will yield valuable information regarding the internal constitution. If the moon contains radioactivity of the same order as the earth, then the elastic energy built up through differential heating and cooling will be released as moonquakes. At present, efforts have been concentrated on obtaining estimates of the eigen frequencies for the free oscillations. The long-period oscillations can be detected by a long-period vertical seismometer and would be especially valuable in yielding information regarding the internal distribution of elasticity.

Computations on the expected spheroidal oscillations of the moon have been carried out by Bolt (1961) and Takeuchi, Saito, and Kobayashi (1961). Takeuchi, Saito, and Kobayashi's results are listed in Table XV where the model of the moon is taken to be homogeneous and having the elasticity of the upper mantle of the earth.

MacDonald (1960) determined the average rate of increase of the principal stress differences for several thermal models of the moon. This increase was on the order of 10 bars per million years. Using this figure, MacDonald estimated the average rate of release of strain energy assuming a finite strength for the lunar materials. The rate of increase of strain energy will be a maximum when the stress differences are on the order of the strength. For a strength of 100 bars, the maximum strain energy per unit volume is on the order of 1.4×10^{3} erg cm^{-3}. The rate of release of strain energy by failure at 100 bars

TABLE XV

Computed Periods of the Spheroidal Oscillations of the Moon
(after Takeuchi, Saito and Kobayashi, 1961)

Order	Period (minutes)	Order	Period (minutes)
2	14.7	6	5.5
3	9.9	7	4.8
4	7.7	8	4.3
5	6.4	9	3.9
		10	3.6

maximum stress difference will be on the order of 4×10^{24} erg year^{-1}. At present, the moon should have a high degree of seismic activity unless the radioactivity is concentrated in a thin layer near the surface or unless the radioactivity of the moon is far less than that of chondritic materials.

The degree of calculated seismic activity varies among models of the distribution of heat sources with depth. In the models studied, the greatest release of strain energy is at a depth of 100 to 700 km. The assumption of deeply buried radioactivity implies a deep foci for lunar seismic disturbances. From such calculations it can be seen that the observation of the average rate of seismic activity on the moon can be used to obtain an estimate of the radioactivity and, possibly, the distribution of radioactivity within the moon.

Thermal Constitution of the Moon

Thermal Radiation from the Moon

Baldwin (1961) has made a study of the thermal radiation of the moon at radio wavelengths in order to limit the possible heat flow at the moon's surface. High-frequency radio and infrared emission determines the temperature at and within a few centimeters of the surface. The longer-wavelength radiation arises from greater depth and, in combination with the observations at higher frequencies, permits a determination of the difference in mean temperature. The uncertainty in the physical constants does not allow an accurate estimate of the depth from which the lower-wavelength radiation is emitted. Baldwin

observed at 168 centimeters and estimated that the depth of penetration is greater than 25 meters.

Baldwin, using various estimates of the thermal conductivity, gives an upper limit to the heat flow of 1 - 10 erg cm^{-2} sec^{-1}. The present rate of heat production in a moon having a chondritic composition would be 3.7×10^{8} erg sec^{-1}, which is equivalent to an equilibrium surface heat flow of 9.6 erg cm^{-2} sec^{-1}. The extreme upper limit of the heat flow obtained from the study of radiowave emission is consistent with the assumption of a chondritic moon; on the whole, a lower heat flow is indicated.

Calculations on the Thermal History of the Moon

Urey (1952, 1957, 1962), Levin (1960), and MacDonald (1959b, 1962) have carried out extensive calculations regarding the distribution of temperature within the moon. The calculations are similar to those described in the preceding sections for the earth. The present thermal state of the moon is determined by the distribution of radioactive elements, the initial temperature, and the manner in which heat is transported. Kopal (1962) argues that convection is required to remove the heat from the earth's interior. Kopal's arguments are based in part on analogy with presumed convection within the earth and, as has been discussed above, the fact that it is unlikely that convection plays an important role in the thermal balance of the earth.

A large number of models of the moon have been constructed in which the radioactivity is uniformly distributed and different thermal

parameters have been assumed (MacDonald, 1960, 1961). Figure 28 shows the result of one such calculation in which the total radioactivity of the moon is equal to that of chondritic meteorites and is uniformly distributed throughout the moon. In addition, the melting points of diopside and iron are shown. As has been discussed earlier, the melting point of a complex multi-component system will certainly lie below the melting point of the pure components. The close approach of the melting curve to the diopside melting point shows that much, if not all, of the moon would be molten at the present time even if the assumed initial temperature is very low.

The fact that the thermal constitution of the moon differs from that of the earth for the same radioactive heat source can be readily understood. The outstanding feature of the temperature-depth curve for models of the moon is the shallow depth at which melting points are reached or exceeded. Since the moon is a relatively small body, the central pressure is low; on the order of 46,000 bars. The low central pressure does not markedly raise the melting point of the possible constituents. The thermal conductivity is sufficiently low so that only the outer portions of the moon are cooling to any extent. The combination of these two circumstances results in the close approach of the melting point curve in the calculated temperature distribution as is illustrated in Figure 28. In the earth, the rate of increase of pressure with depth is much larger, and raises the melting point with the result that the earth's mantle is solid.

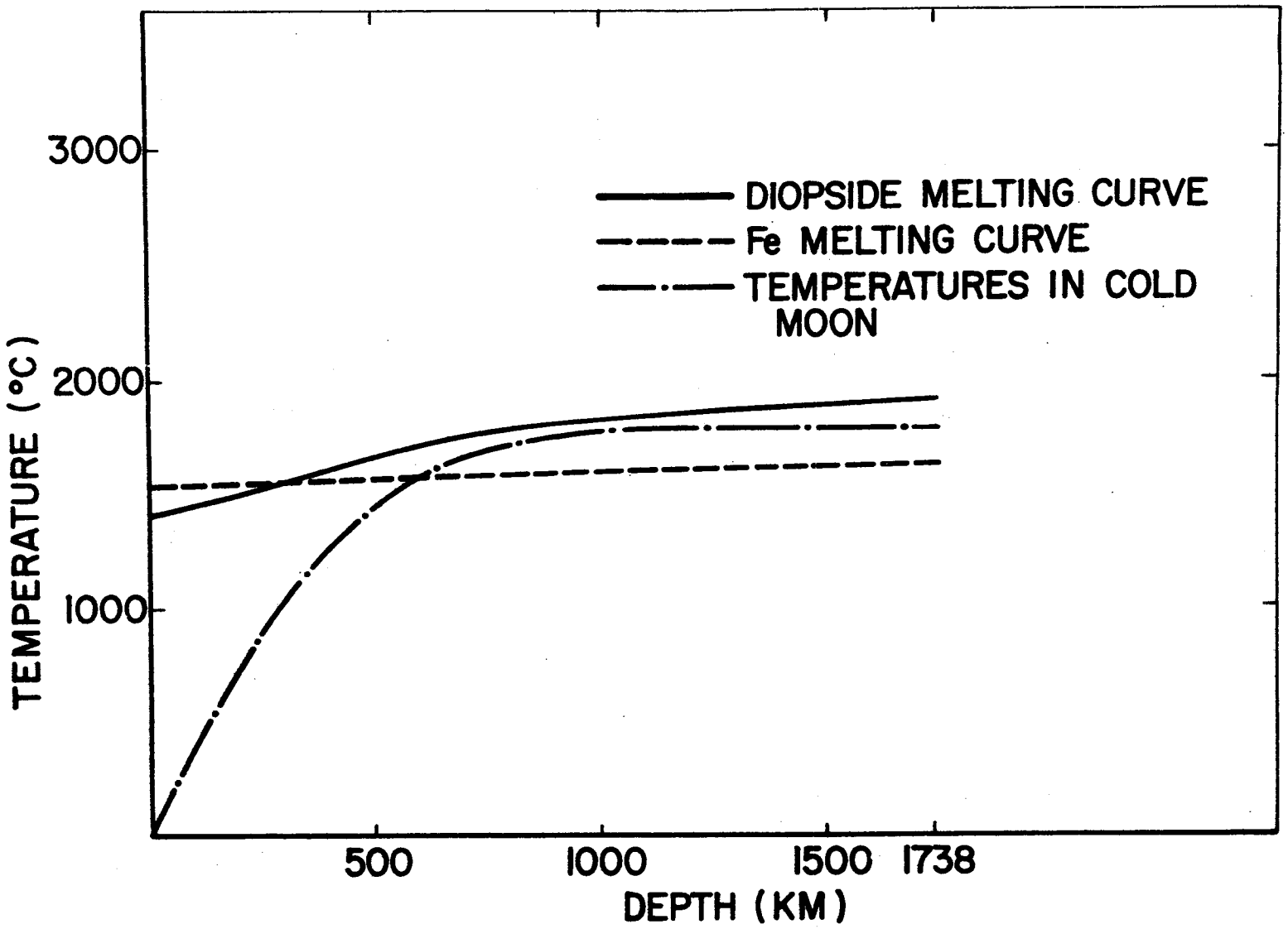


Figure 28: The variation of temperature in a moon having the chondritic radioactivity uniformly distributed. Initial temperature is $0^{\circ}C$.

In a homogeneous moon, the melting temperature of silicates is approached or exceeded provided the moon has a composition of chondritic meteorites. The assumptions of chondritic radioactivity and inhomogeneity are then inconsistent with the astronomical data on the figure of the moon provided these data are interpreted in terms of internal strength. There are two possible solutions:

1) The total radioactivity of the moon is less than that of chondritic meteorites.

2) The moon is a differentiated body in which the radioactivity is concentrated toward the surface. (See Figure 29)

Both of these alternatives have been noticed by Urey (1957) and he has carried out calculations on the effect of a lower radioactivity.

Figure 30 shows the effect of lowering the total radioactivity on the maximum temperature. The radioactivity is assumed uniformly distributed. In this case, the maximum temperature is reached at the center of the moon. The central temperature is plotted as a function of the fraction of chondritic radioactivity maintaining the chondritic ratio among the heat-producing elements. If the initial temperature is zero, the central temperature lies below the melting temperature for a total radioactivity equal to that of chondritic meteorites. An initial temperature of 1200°C requires that the radioactivity be less than 0.4 that of chondritic meteorites for the present central temperature to lie below the melting point of diopside. As has been noted above, this is an upper limit to the possible internal temperature and the actual concentration of radioactivity would have to be on the order of 0.3 that of chondritic meteorites.

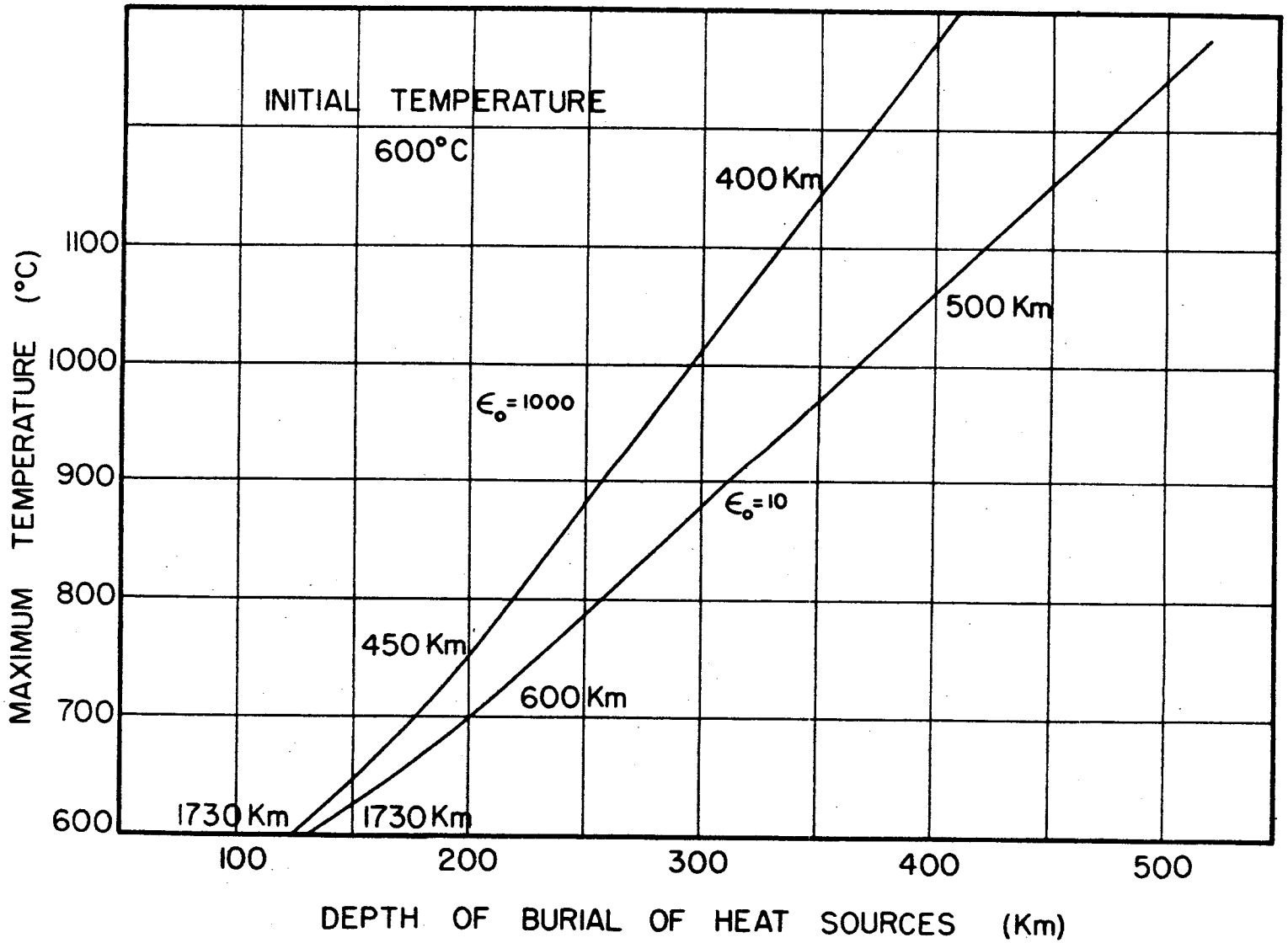


Figure 29: Maximum temperature within the moon as a function of the depth of burial of heat sources. Depth at which maximum temperature is reached is shown for various depths of burial. Initial temperature is 600°C and total radioactivity equals that of chondrites.

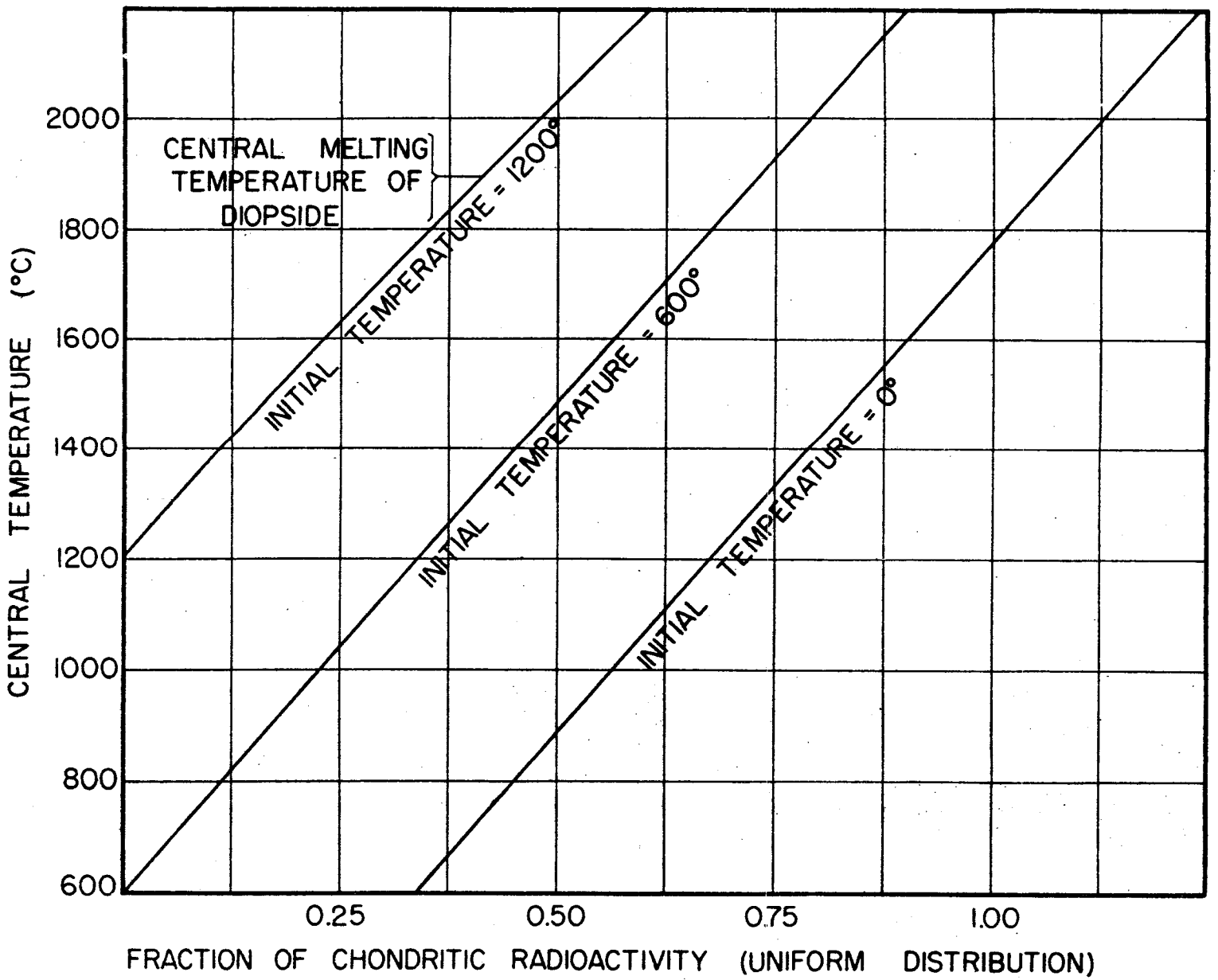


Figure 30: Maximum temperature within the moon as a function of the total radioactivity of the moon. Radioactivity is uniformly distributed.

An experiment of first importance in the investigation of the moon's interior is the determination of the outflow of heat. Figure 31 shows the surface heat flow as a function of the depth of burial of heat sources, assuming chondritic radioactivity and an initial temperature of 600°C. Two values for the opacity are shown and these indicate that the detailed assumptions regarding the thermal conductivity are not of great importance in determining the surface heat flow. If the heat sources extend to a depth of some 500 km, then the surface heat flow is twice that of a model in which the heat sources are located within the upper 70 km. The near-surface concentration of heat sources permits heat produced during the early stages of the history of the moon to escape rather than to be trapped. If the heat sources are buried to a depth of some 500 km, most of the heat is reaching the surface.

The dependence of surface heat flow on the amount of radioactivity is illustrated in Figure 32 for two initial temperature distributions. In these models, the radioactivity is assumed to be uniformly distributed throughout the moon.

The hypothesis that the moon is differentiated provides a solution to the discrepancy between its figure and the inferred high temperature of a moon of chondritic composition. O'Keefe and Cameron (1962) argue that the surface features of the moon do indeed indicate a differentiated body, since the irregularities in the topography are in isostatic equilibrium. Kopal postulates that the excess heat is removed by convection currents.

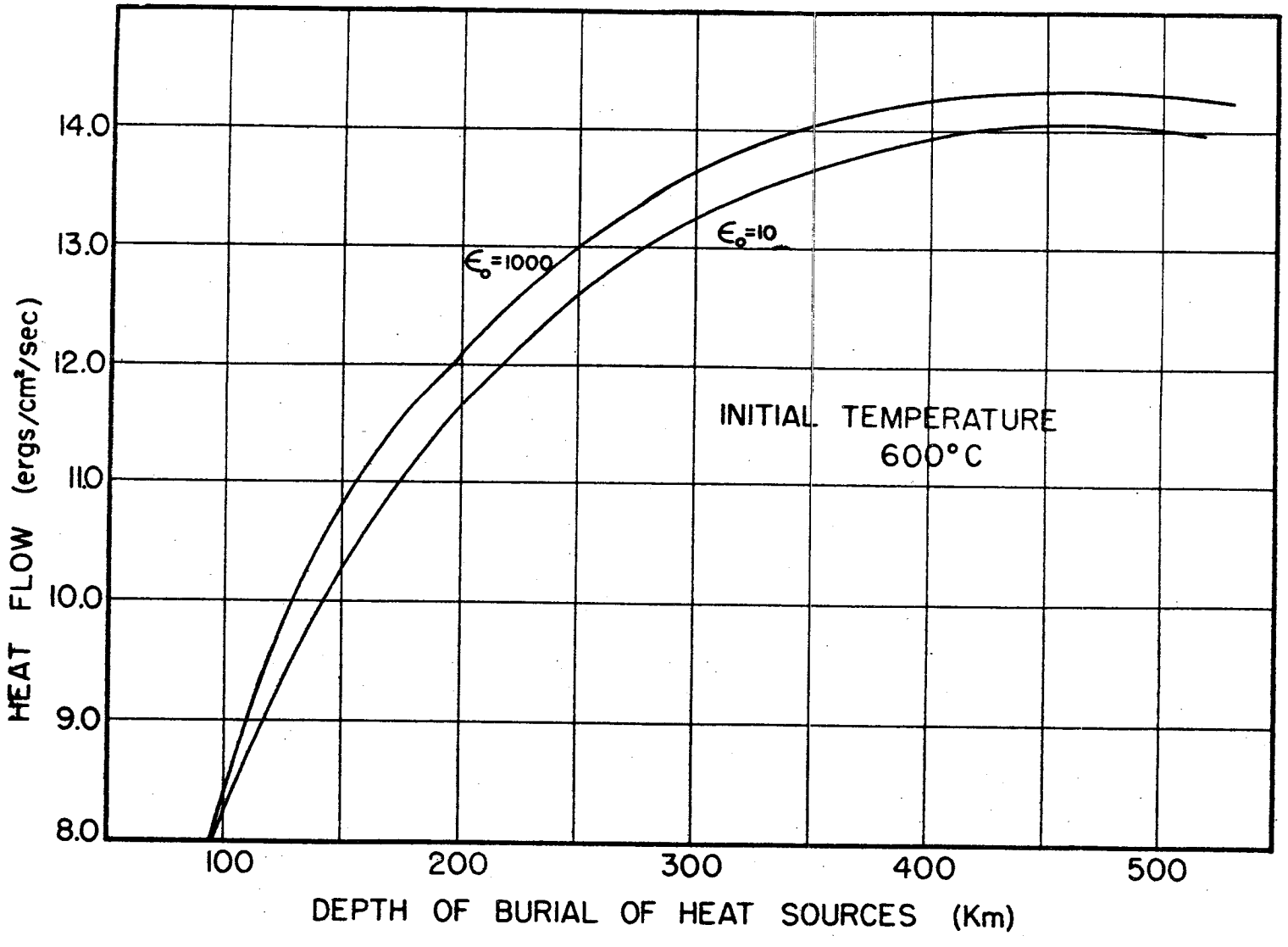


Figure 31: Dependence of surface heat flow on the depth of burial of heat sources. Total radioactivity equals that of chondrites. Initial temperature is 600°C.

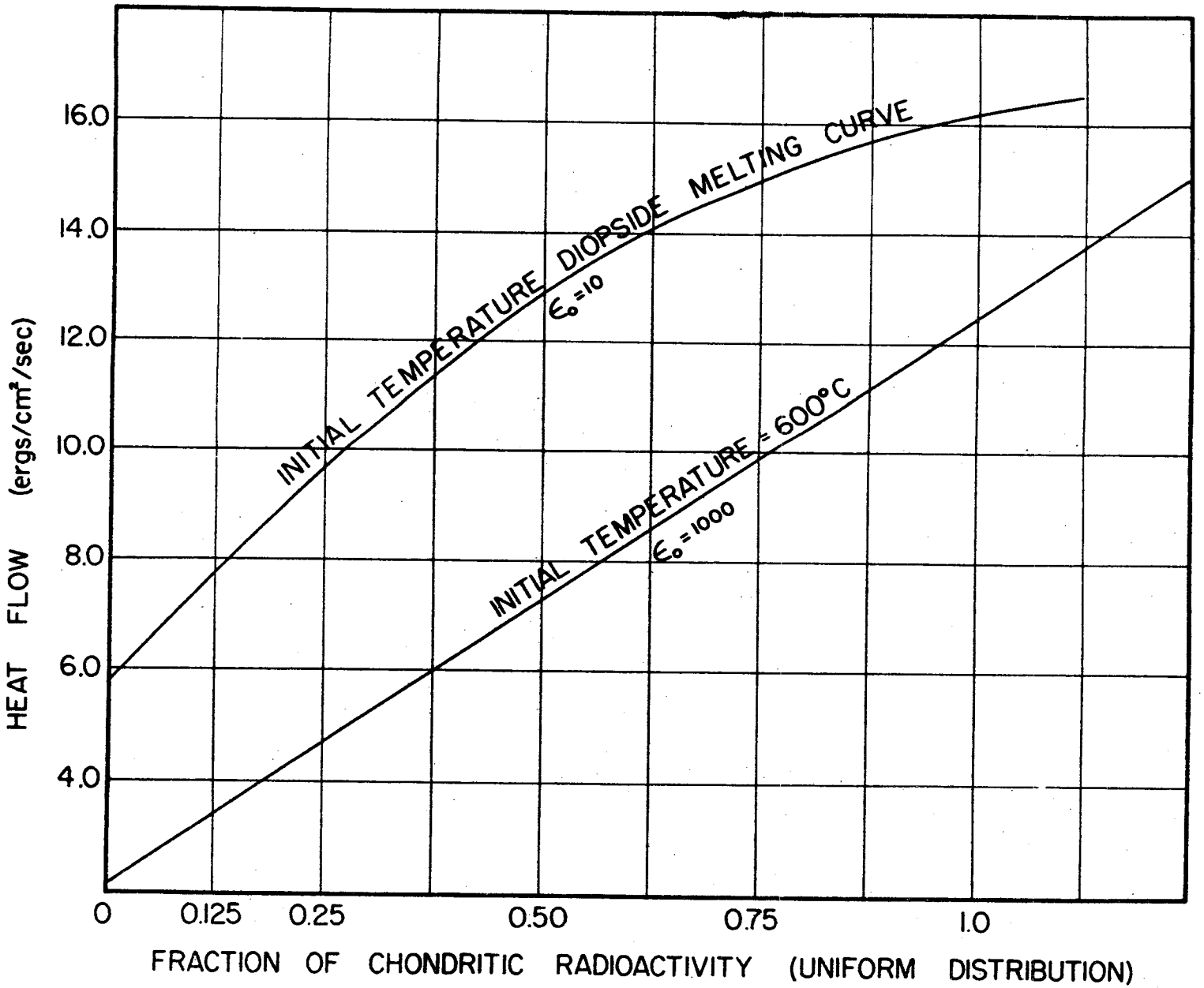


Figure 32: The variation of surface heat flow with total radioactivity.

The principal argument against large-scale convection within the moon is based on the absence of faults showing a transverse displacement. The moon's surface is covered by circular craters and linear features. As yet, no evidence has been produced that there has been large-scale horizontal displacement of these features. As a result, it is unlikely that convection plays a major role in establishing the thermal field within the moon.

An alternative supposition is that the radioactive element composition of the moon differs from that of the chondritic meteorites. This hypothesis is strengthened by the fact that the density of the moon is appreciably less than the density of chondritic material.

Much of the uncertainty concerning the moon's interior is a reflection of the uncertainty regarding the interpretation of the surface features of the moon. Baldwin (1962) has presented a detailed review of current thoughts regarding the nature of the lunar surface. Most workers would agree that many of the superficial features are due to meteorite bombardment. There remains the question of the nature of the surface that was bombarded by the meteorites. A possible answer is provided in calculations carried out by MacDonald (1960). A moon of chondritic composition would initially expand at a rapid rate, with the radius changing some 3 km during the first 10^9 years. During this epoch, tension fractures would form and material at depth would melt provided that the accumulation process left the moon at a temperature on the order of 600° to 1000°C. The partial melting of chondritic

material would produce magma of basaltic composition which would rise through the fractures to the surface. In this way, a crust of a few kilometers of basalt could form while the bulk of the interior would never be completely molten. Thus, if the moon does indeed have as high a concentration of radioactivity as that of chondritic meteorites, then one would expect a crust of a few kilometers composed of basaltic material overlying more mafic rocks with a composition intermediate between basalt and dunite. Granite or other siliceous rocks would not be expected since, on the earth, these rocks are the result of a long sequence of processes involving erosion, sedimentation, and remelting associated with continent formation.

The Moon's Magnetic Field

The flight of Lunik II demonstrated that the surface field of the moon on the sunlit hemisphere was less than 10^{-3} gauss (Dolginov, et al., 1961). This observation is consistent with current views regarding both the origin of planetary magnetic fields and the internal constitution of the moon. The low density of the moon implies that the moon cannot have a metallic core of significant dimension. Since current theories of planetary magnetism require a fluid, conducting core, there is no reason to expect that the moon should have a magnetic field.

IV. INTERNAL STRUCTURE OF MARS

Introduction

The mass of Mars is well determined by the orbital constants of the satellites, Phobos and Deimos. The density is uncertain because the diameter of the solid is disputed. The orbital constants for Phobos and Deimos also provide an accurate value of $\frac{C - A}{M_{\sigma} R_{\sigma}^2}$. The ratio can be converted into an estimate of the moment of inertia C provided hydrostatic equilibrium is assumed. The problem is then to construct a model of the density distribution within the planet such that the variation of density yields the observed mass and moment of inertia. Without a definitive value for the radius, the problem is open and there are a number of possible models of the density distribution within Mars. In particular, it is not clear whether or not we should expect Mars to have an inner core.

Geometric Figure of Mars

Urey (1952) has discussed the difficult problem of obtaining the radius of the solid surface. Rabe (1926) obtained a value of 3415 km. This value is usually interpreted as the correct visual radius. Later workers showed that Mars had different diameters for light at different wavelengths. Trumpler (1926) found that the diameter in yellow light is 3310 km and obtained a value agreeing with that of Rabe for white light. Dollfus (1952) made another attempt and arrived at a value 25 km greater for the radius. Camichel (1954), in a detailed study, obtained results which appear to confirm Trumpler's value for the radius of the solid surface. Recently, Dollfus (1962) has again investigated

the problem of the geometric figure. He obtains a mean radius of 3382 km, but there is an uncertainty on the order of 30 km due to the difficulties in interpreting the effects of the atmosphere. Table XVI summarizes the radius and mean density for Mars obtained by the principal investigators.

Optical observations on the disk show a consistent difference between the polar and equatorial diameter. This difference appears to be greater than the errors in the observations. Dollfus (1962) obtained a difference in polar and equatorial radius of 40 km which corresponds to an optical flattening of 0.0017. Whether this optical flattening refers to the solid surface or not is a matter of dispute. As we shall see, it is impossible to interpret this flattening in terms of the internal constitution.

Gravitational Field

Using Dollfus' value for the radius and Woolard's (1944) study of the motion of Phobos, we obtain $J_2 = \frac{C - A}{M_\sigma R_\sigma^2} = 0.00197$, where R is the equatorial radius. This value is observed and does not depend on any assumptions about the attainment of hydrostatic equilibrium within the planet. Assuming that the planet is indeed in hydrostatic equilibrium, it is then possible to calculate a flattening of 0.0051 (see Equation 8), and from the flattening obtain a value for the polar moment of inertia (see Equation 7). Both the dynamical flattening and the moment of inertia depend on the assumption that the planet is in hydrostatic equilibrium.

TABLE XVI

Radius and Mean Density of Mars

Reference	Radius (km)	Mean Density g cm^{-3}
Rabe (1926)	3415	3.84
Trumpler (1926)	3310	4.21
Dollfus (1952)	3335	4.12
Dollfus (1962)	3382	3.98

The flattening obtained from dynamical considerations is less than half the flattening derived from optical observations. The explanation of this difference is critical to the understanding of the internal constitution of Mars. We first consider whether the discrepancy could arise from the internal strength of the planet. The critical quantity in the argument is the maximum stress difference which the planet Mars can support. The magnitude of the irregularities in the gravitational field will always be limited by the strength of the material which supports the implied density inhomogeneities. The dimensions of stress (force per unit area) are $ML^{-1}T^{-2}$. Since the gravitational constant (dimension $L^3M^{-1}T^{-2}$) is the same everywhere, time must scale between two planets as $M^{-1/2}L^{3/2}$. Substituting this time-scaling in the stress-scaling yields M^2L^{-4} as the scaling law for the planet. For a given strength, Mars will support an inequality in figure $(M_\oplus^2/R_\oplus^4)/(M_\sigma^2/R_\sigma^4)$ times as large as the earth. Substituting numerical values, we show that Mars will support inequalities in figure 6.8 times as large as that on the earth provided that the basic strength of the materials within Mars is the same. The earth supports a deviation ΔJ_2 that is 1.11 per cent of the equilibrium J_2. For Mars, this implies that the internal strength could support a flattening as large as 0.0053. Even if the material within Mars were ten times as strong as that of the earth and thus capable of supporting stress differences on the order of 20,000 bars, the maximum flattening that could be supported would be about 0.007. All these values are substantially lower than the observed flattening of 0.0117. A major mystery is thus the origin of the optically observed equatorial bulge.

Density Distribution within Mars

The construction of a density model for Mars requires the assumption of a surface density and a law of the variation of density with pressure. In order to meet the requirements provided by the total mass and the moment of inertia calculated on the basis of hydrostatic equilibrium, an additional degree of freedom is required. Jeffreys (1937) suggests a chemically distinct core in analogy to the structure of the earth. The free parameter to be determined for a given surface density is the core radius. In addition, the laboratory data and an interpretation of the structure of the earth's mantle suggests that phase transitions and major changes in density take place in silicates at pressures on the order of $1.0 \times 10^5 - 1.5 \times 10^5$ bars.

MacDonald (1962) has carried out detailed calculations on the possible density distributions for Mars, assuming that the equation of state for the material is that followed by materials in the earth's mantle. The nature of the calculations is illustrated in Figure 33 in which consistent values of the core radius and surface density are shown as a function of the flattening. If the planet is in hydrostatic equilibrium, the flattening is 0.0051 and the planet could support a core of 1180 km radius provided the surface density were 3.58. Similar calculations have been carried out for other values of the radius for Mars. If the lower values of the radius for Mars are adopted, it would appear that the only possible model consistent with the dynamically obtained value of J_2 is the one in which the surface density is large, about 3.8 - 4.0 and where there is no large-scale chemical

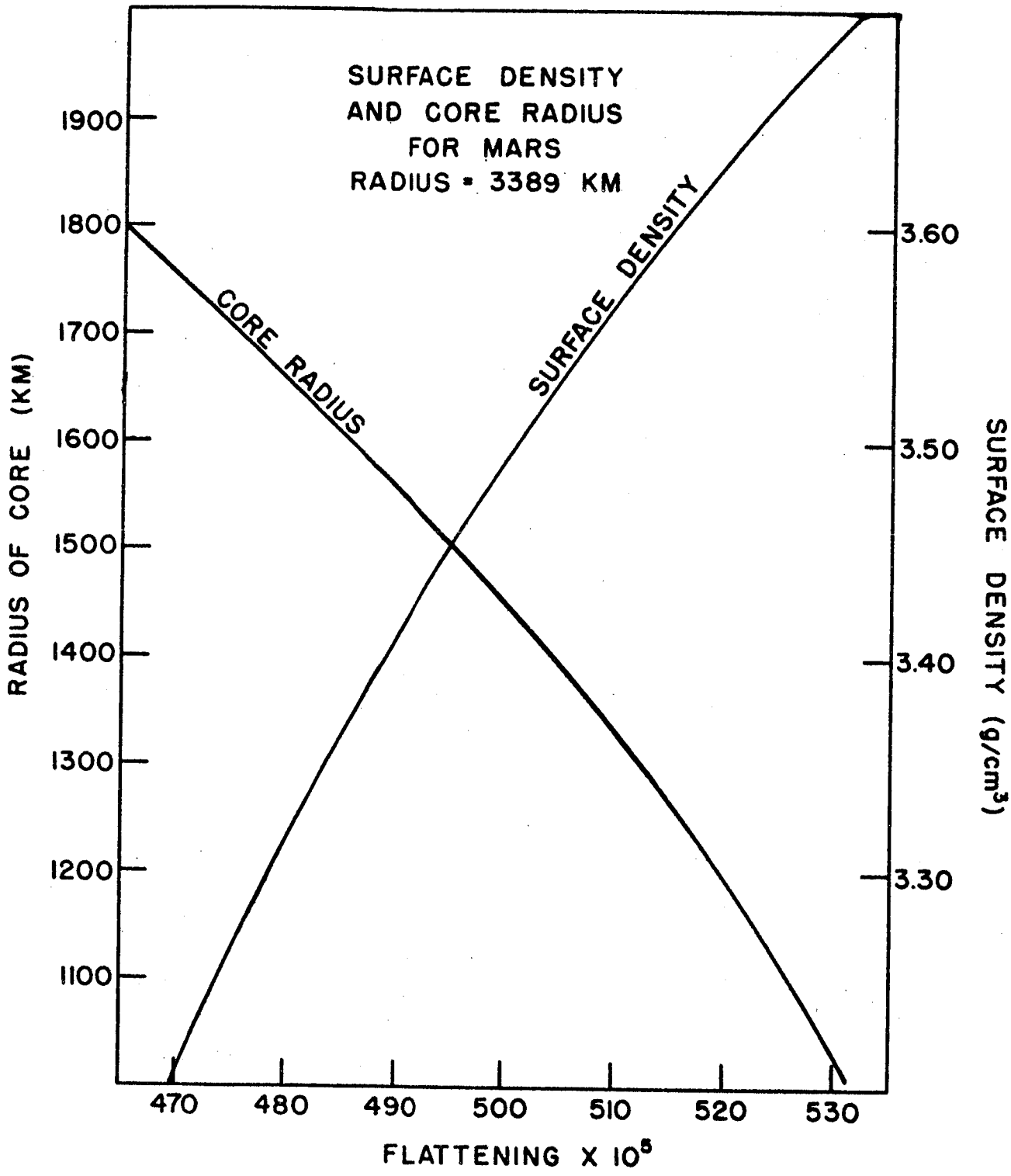


Figure 33: Dependence of surface density and radius of the inner core of Mars as a function of the hydrostatic flattening.

inhomogeneity in the form of an inner metallic core. The difficulty in accepting such a model is that the surface density is large compared with the density of the moon and chondrites.

Figures 34 and 35 illustrate possible density models for Mars. In both cases the possibility of a phase transition at pressures of 1.0×10^5 bars has been taken into account. The effect of the phase transition places dense material toward the center. In both cases only a small core is allowed.

From the studies of possible density models, MacDonald (1962) concludes that Mars has a structure intermediate between that of the moon and that of the earth. Differentiation has proceeded sufficiently to form a small metallic core, but the proportion of the mass within the core is less than in the earth. The indicated high surface densities suggest that the metallic phase is still distributed throughout the outer regions. Because of the low melting point of the metallic phase, this places a restriction on the possible temperature distributions within the planets.

Thermal Constitution of Mars

MacDonald (1962) carried out detailed calculations on the possible temperature distribution within Mars assuming that the composition is that of chondritic meteorites. These calculations are shown in Figures 36 and 37 for the parameters given in Table XVII. The outstanding feature of the temperature distribution is that if a chondritic

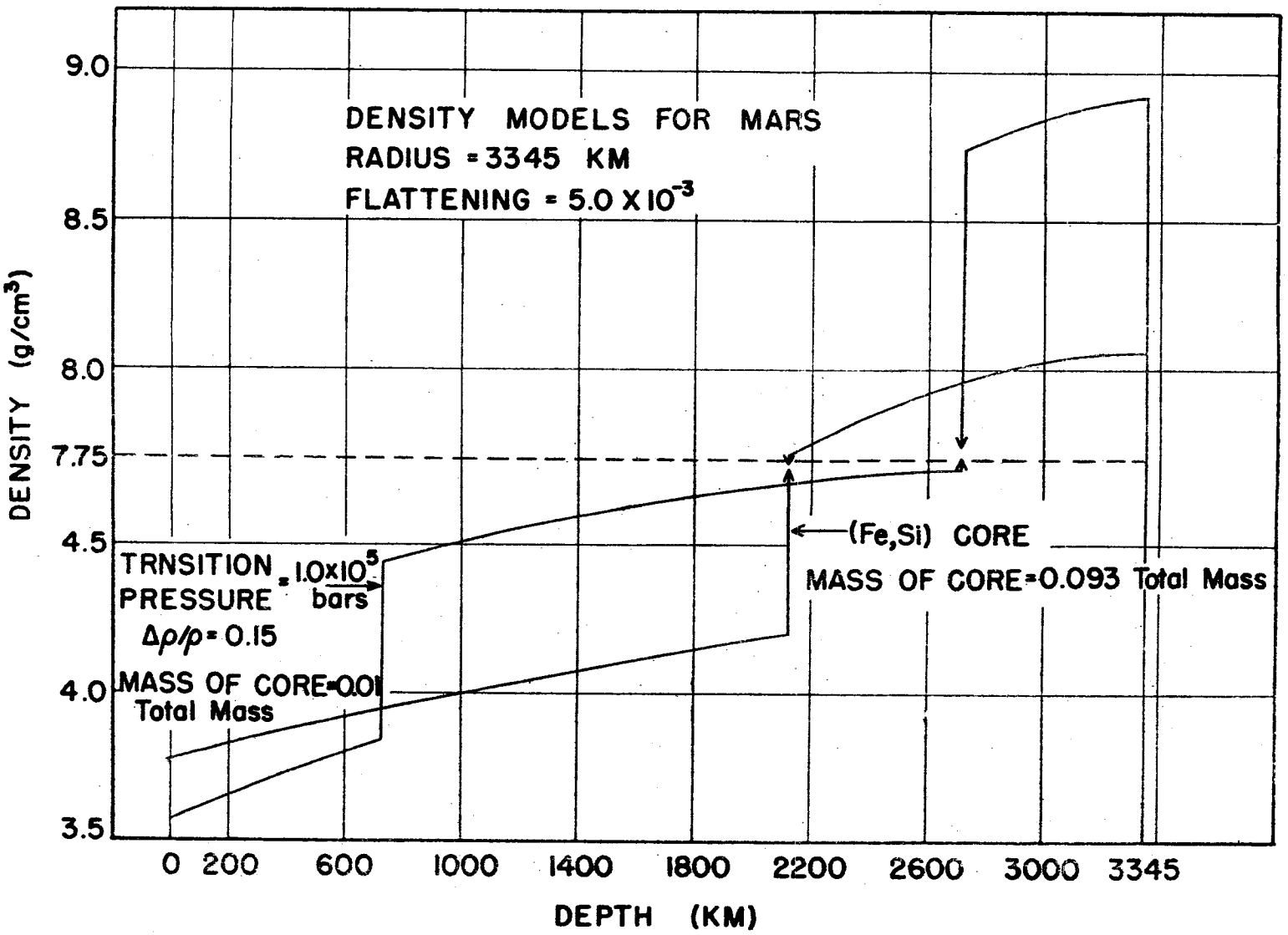


Figure 34: The variation of density with depth in Mars for two models. Outer radius = 3345 km, flattening = 5×10^{-3}. In one model, the phase transition takes place at 10^5 bars and there is a relative change in volume of 15 per cent. In the other model, no phase transition is assumed. The mass of the chemically distinct core is 0.01 of the total mass in the phase transition model and 0.093 of the total mass in the model with no phase transition.

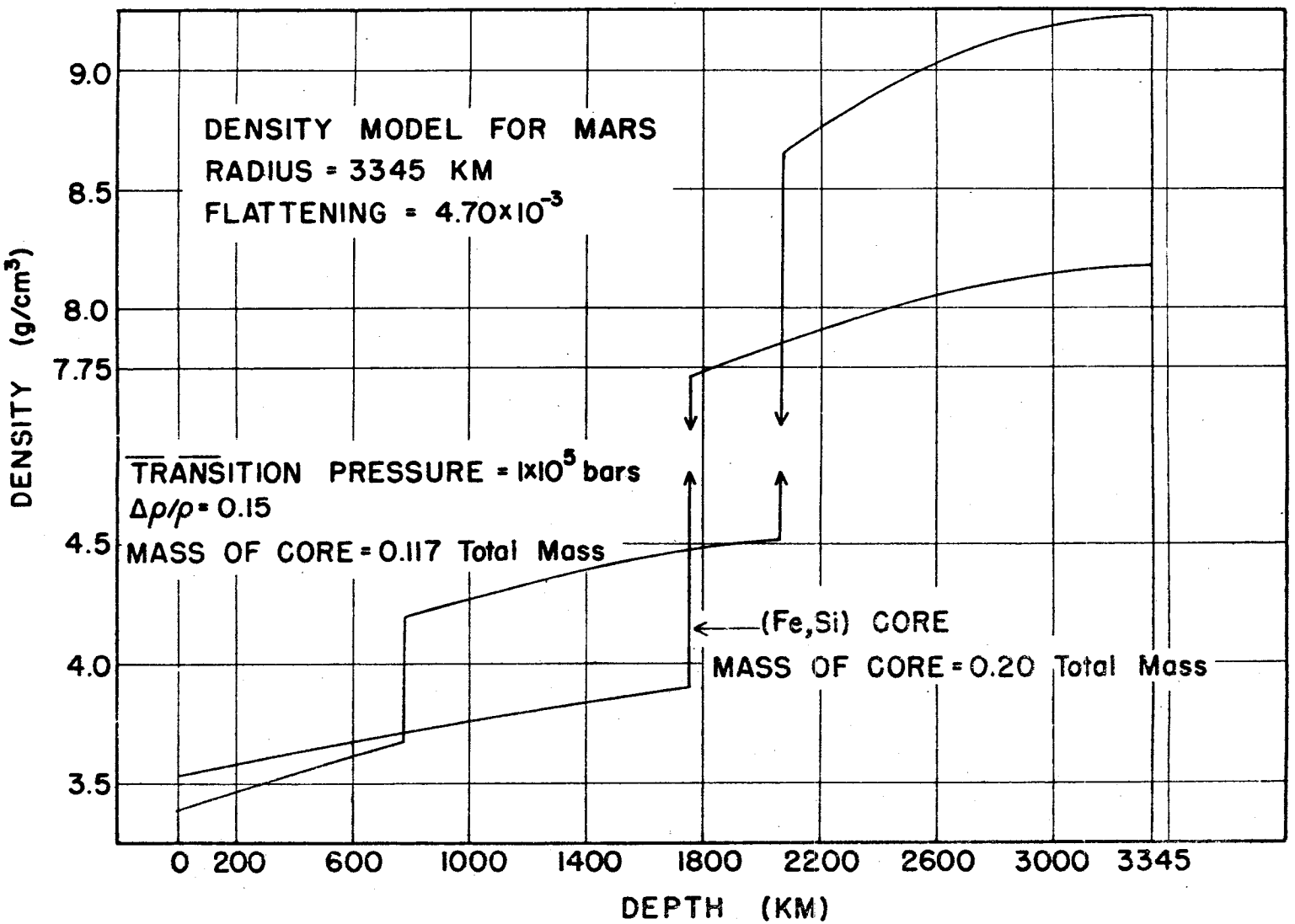


Figure 35: Density models for Mars with outer radius of 3345 km and a flattening of 4.7×10^{-3}. The phase transition model assumes a transition pressure at 1×10^5 bars and a relative change of volume of 15 per cent. In this model, the mass of the core equals 0.117 of the total mass. In the model with no phase transition, the mass of the core equals 0.20 of the total mass.

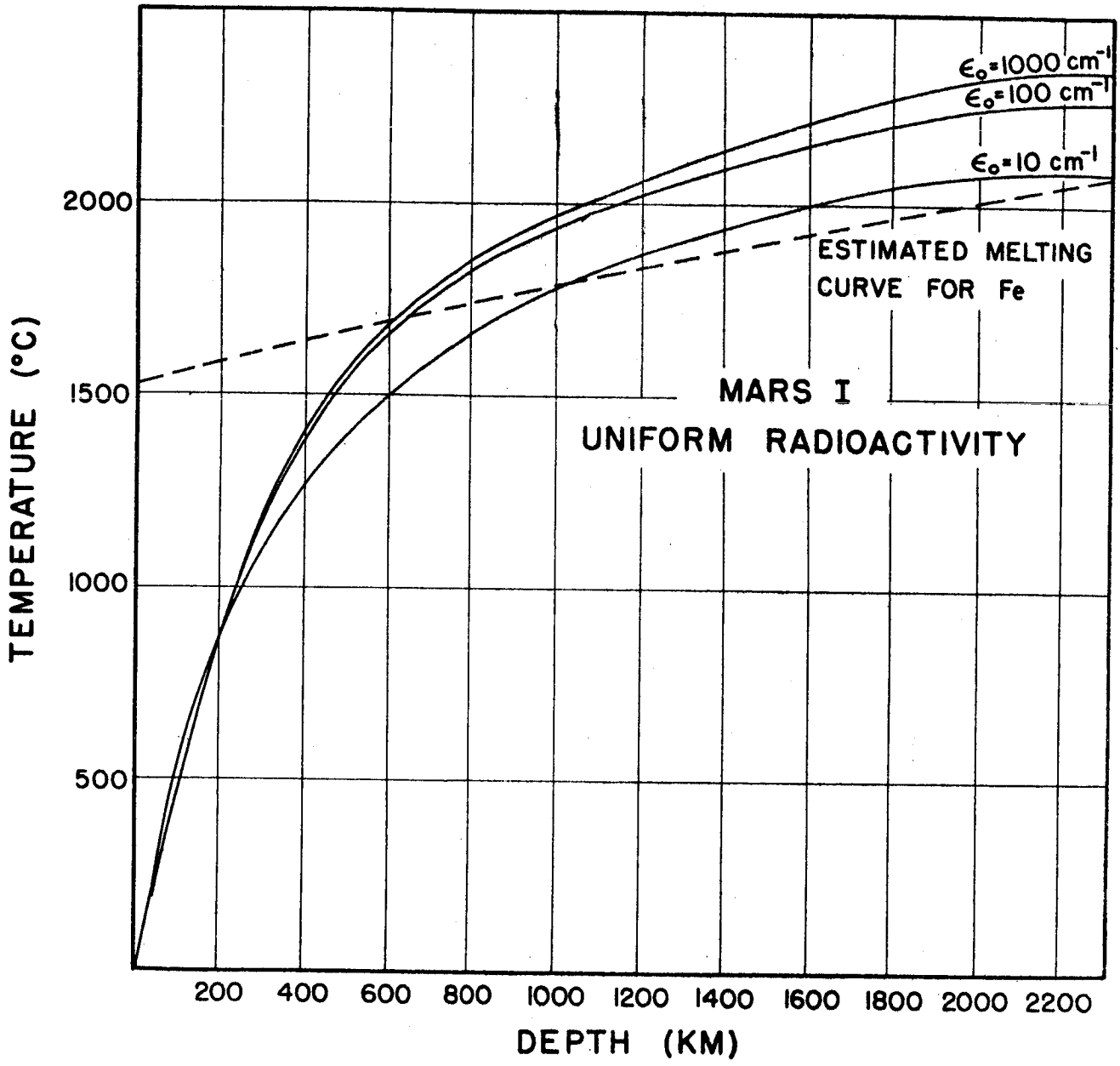


Figure 36: Calculated temperature distribution of Mars assuming uniform chondritic radioactivity. The detailed parameters are given in Table XVII.

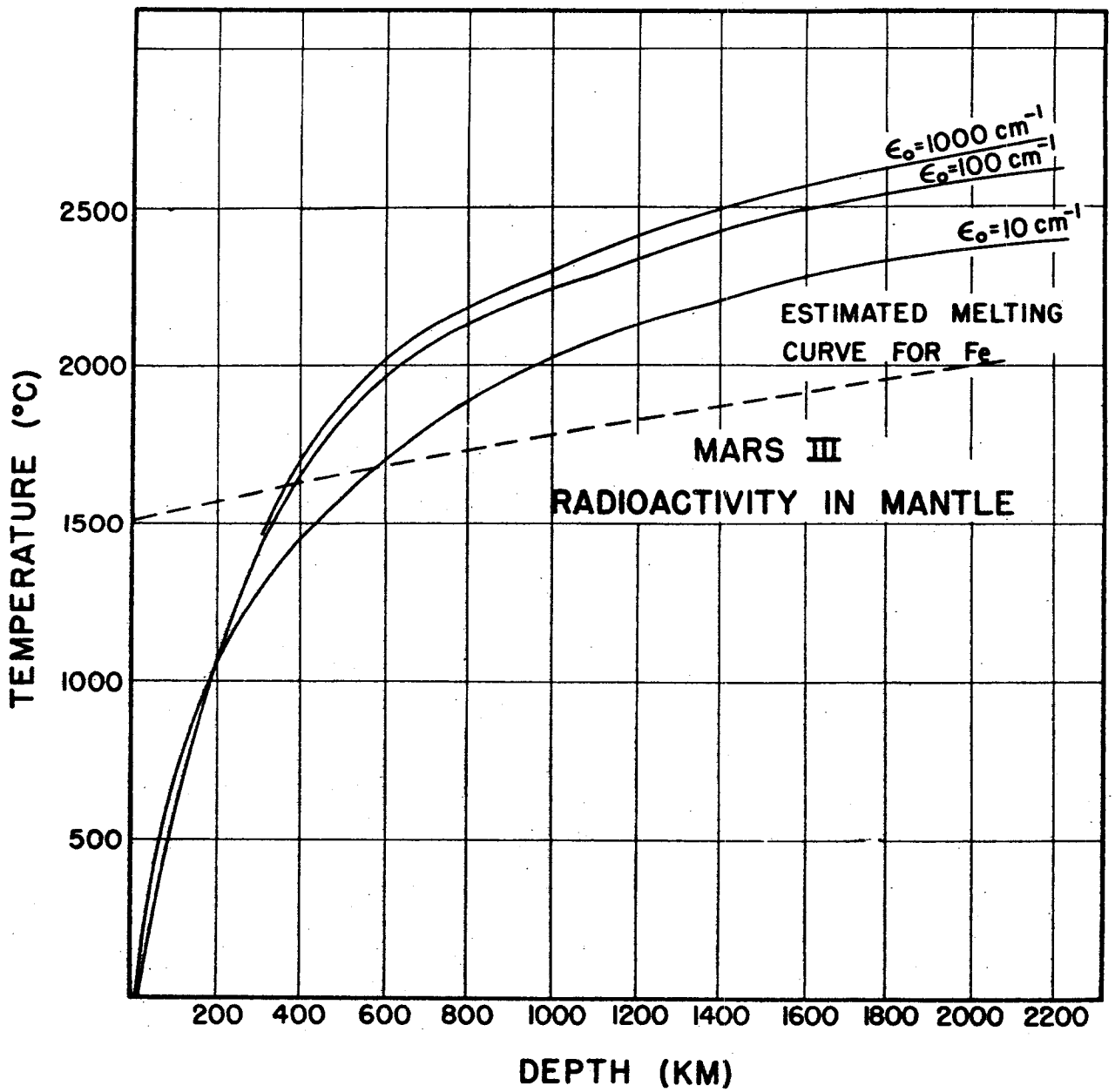


Figure 37: Calculated thermal distribution in Mars with radioactivity concentrated in outer two-thirds of the planet.

TABLE XVII

Thermal Parameters for Mars Models I and III

Depth km	Initial Temperature ^{o}C
0	-39
330	100
1040	400
1750	700
2000	800
2300	910
Radius	3313 km
Surface Temperature (constant in time)	$-39^{o}C$

Radioactive Element Concentration ($g\ g^{-1}$)	Mars I	Mars III
	$U = 1.1 \times 10^{-8}$	1.35×10^{-8}
	$Th = 4.4 \times 10^{-8}$	5.40×10^{-8}
	$K = 8.0 \times 10^{-4}$	9.8×10^{-4}
Time of integration	4.51×10^{9} years	
Lattice conductivity	2.5×10^{5} erg $cm^{-1}\ ^{o}C^{-1}$	
Density (uniform)	3.9 $g\ cm^{-3}$	
Heat capacity	1.46×10^{7} ergs $g^{-1}\ ^{o}C^{-1}$	

Final Surface Heat Flow erg $cm^{-2} sec^{-1}$	Mars I	Mars III
ϵ_o = for 10 cm^{-1}	17.0	19.7
ϵ_o = 100 cm^{-1}	12.5	16.0
ϵ_o = 1000 cm^{-1}	12.0	15.0

radioactivity is assumed, then the temperature exceeds the melting temperature of iron in shallow depths. Chondrites contain 12 per cent metallic iron. If Mars were indeed of chondritic composition, the radioactivity would lead to the melting of the iron, the iron would gravitationally differentiate to form a core composing about 10 per cent of the total mass of the planet. The resulting surface density would be too low to account for the observed mass and moment of inertia. It would thus appear that the concentration of radioactive elements within Mars must be less than that found in the chondritic meteorites. The solution to the problem of the internal constitution of Mars awaits a definitive value for the radius of the solid body and a better understanding of the discrepancy between the dynamical and optical flattening of the body.

Magnetic Field of Mars

The calculations discussed above suggest that Mars contains a small core containing about 10 per cent of the total mass of the planet. This metallic core could provide a magnetic field. If planetary magnetic fields result from thermally-driven convection, then Mars would be expected to have a magnetic field. The relatively high rate of rotation of Mars would ensure that the field would be primarily dipolar, oriented along the axis of rotation as in the case of the earth. Alternatively, if planetary fields are driven by precessional torques, then Mars would not be expected to have a field comparable in strength to that of the earth. The torques exerted by the two small satellites are insufficient to transfer the needed momentum.

V. THE INTERNAL STRUCTURE OF VENUS AND MERCURY AND THE ROTATION OF THE PLANETS

Introduction

Little can be deduced about the internal structure of Venus and Mercury. Smith, et al., (1963) have shown that the magnetic dipole moment of Venus is at most 1/10 to 1/20 that of the earth. This can either be interpreted in terms of the internal constitution of the planet or in terms of the theory of the origin of the magnetic field. Mercury has an anomalously high density, the origin of which is not understood.

Both Mercury and Venus rotate at a rate comparable to their orbital periods. This low rate of rotation may indirectly provide information about the thermal conditions within the planet. The argument is based on the observation that the rates of rotation of six of the planets are remarkably similar despite major differences in orbital characteristics, masses, and moments of inertia. This is illustrated in Figure 38, where the density of angular momentum is shown as a function of mass for the planets. Venus and Mercury show a large deficiency in angular momentum compared to the major planets and Mars. The earth also shows a deficiency in angular momentum for its mass. The earth's deficiency can be understood in terms of the tidal deceleration which the earth has experienced due to interaction with the moon. The uncertainties in angular momentum indicated in Figure 38 reflect both the uncertainties in the angular velocity and in the moment of inertia. Recent radar studies (Pettengill, et al.,

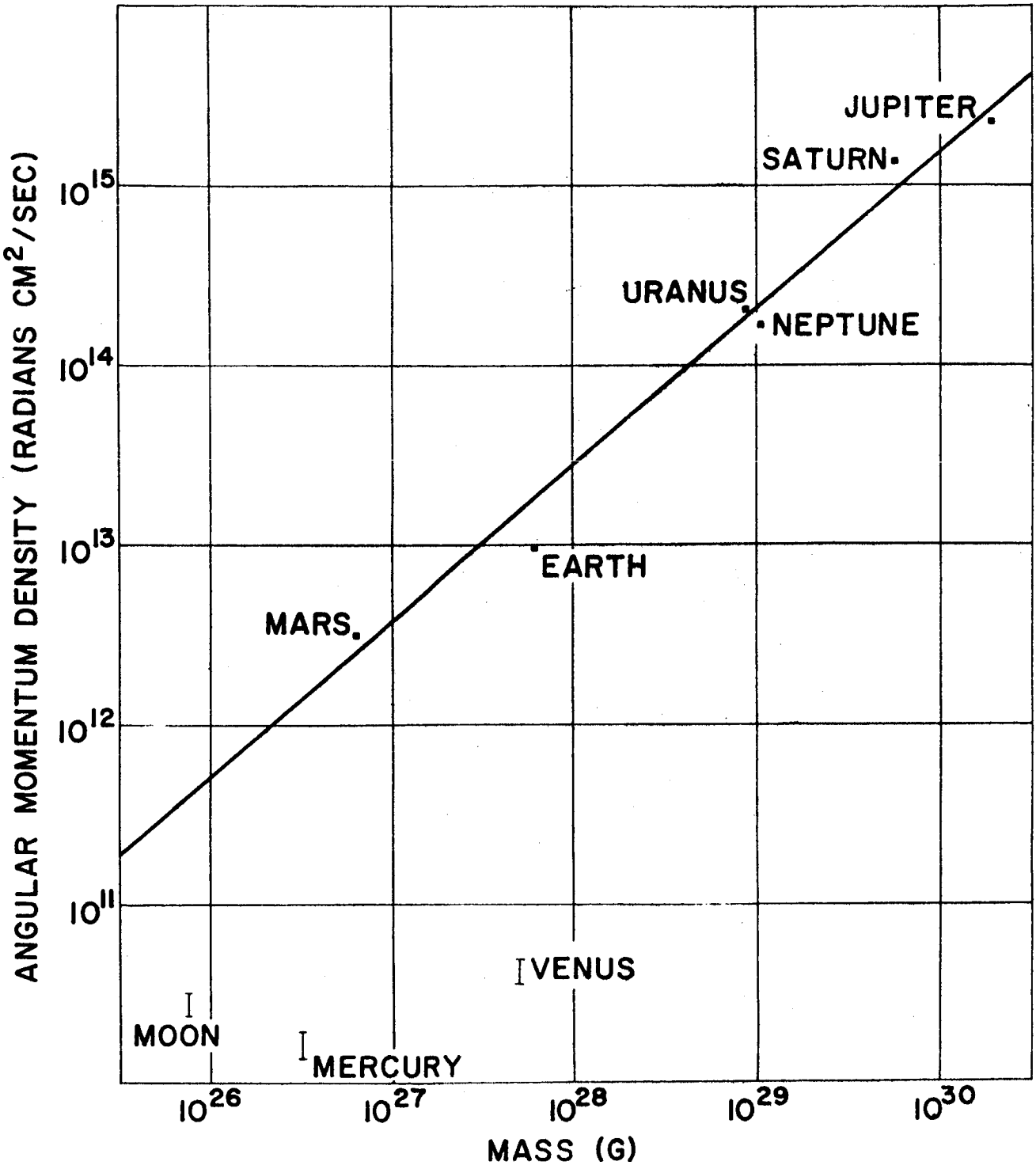


Figure 38: The density of angular momentum associated with the rotation for the planets. The vertical bars reflect the uncertainties in the angular velocities and moments of inertia of the moon, Venus, and Mercury.

1962; Goldstein and Carpenter, 1963) show that Venus probably rotates in a direction opposite to that of the earth at the rate of only one revolution in 240 days. The rate of rotation of Mercury is uncertain, but what evidence there is points to a period about equal to the orbital period (Russell, Dugan, and Stewart, 1945).

Rotation of Venus and Mercury

Let us suppose that the initial angular momentum of Mercury and Venus is appropriate to the mass as given by the straight line in Figure 38. In this hypothesis the present angular velocity of Mercury and Venus are the result of tidal interaction with the sun. Jeffreys (1959) shows that the rate of change in the angular velocity of rotation is proportional to the sixth power of the apparent diameter of the planet as seen from the sun and to sin ϵ where ϵ is the phase lag of the solar tide. The rate of change of angular velocity of Venus due to solar tidal interaction is about 7 times as great as that of the solar retardation of the earth provided the phase lag on Venus equals the phase lag on the earth (see Table V). With equal phase lags, the present rate of deceleration of Venus is nearly equal to the deceleration of the earth under the combined influences of the moon and the sun. In the past, the tidal deceleration of the earth was greater because of the closer approach of the moon. If Venus indeed had an initial angular momentum density on the order of 10^{13} cm^2 sec^{-1}, then its present angular velocity can be accounted for only if the rate of tidal dissipation is greater within Venus than within the earth. In the earth, the tidal dissipation takes place both within the body of

the earth and in the oceans (Munk and MacDonald, 1960b). The solid dissipation is due mainly to internal friction (Knopoff and MacDonald, 1958). At temperatures near or at the melting temperature, other mechanisms (creep, elastoviscosity) become dominant. If there are large regions within the interior of Venus near or at the melting point, then tidal dissipation would be greater than on the earth. A mixture of molten and solid silicates would be particularly effective in dissipating energy since the energy loss would be through direct viscous interaction

Thermal Structure of Venus

The observed angular velocity of Venus may be explained by supposing that the temperatures at depth within the planet are closer to the melting point than the temperatures within the earth. It appears likely that in the earth the temperature comes closest to the melting temperature at depths on the order of 100 to 200 km, but below this depth the temperatures may be well below the fusion temperature. The pressure at a given depth in Venus is probably only a few per cent less than the pressure at the same depth within the earth. The melting temperatures of silicates at a given depth in Venus will be approximately equal to the corresponding melting temperatures within the earth. If the concentration of heat-producing radioactive elements within Venus is similar to that in the earth, the actual temperature at a given depth in Venus should be higher than the

earth. Extensive observations of microwave emission spectrum from Venus show that the spectrum is flat between 3 and 21 centimeters and corresponds to a brightness temperature of 600°K (Mayer, 1961). These ground-based observations have been substantiated by the flight of Mariner II (Barath, et al., 1963). If this high surface temperature has been maintained for a reasonable fraction of the history of the planet, then the internal thermal conditions will be affected in a major way. The higher surface temperature on Venus implies that at the same depth the temperature in that planet will be higher than in the earth.

The possible thermal structure within Venus is illustrated in Figures 39 and 40. Basic data for these calculations are given in Table XVIII. The principal result of the calculations is that for a surface temperature of 600°K the temperature at a depth of 200 km is some 200° higher than in a model in which the temperature is 300°K.

The low rate of rotation of Venus is compatible with a model in which a substantial fraction of the planet is at or near the melting point. The high surface temperature combined with a radioactive composition equal to that of chondritic meteorites would provide a consistent model. The evidence is sketchy, but present-day data are consistent with the hypothesis that the radioactive chemical composition of Venus and the earth are similar, but that the high surface temperature of Venus has raised the temperature at depth above the melting point of silicates.

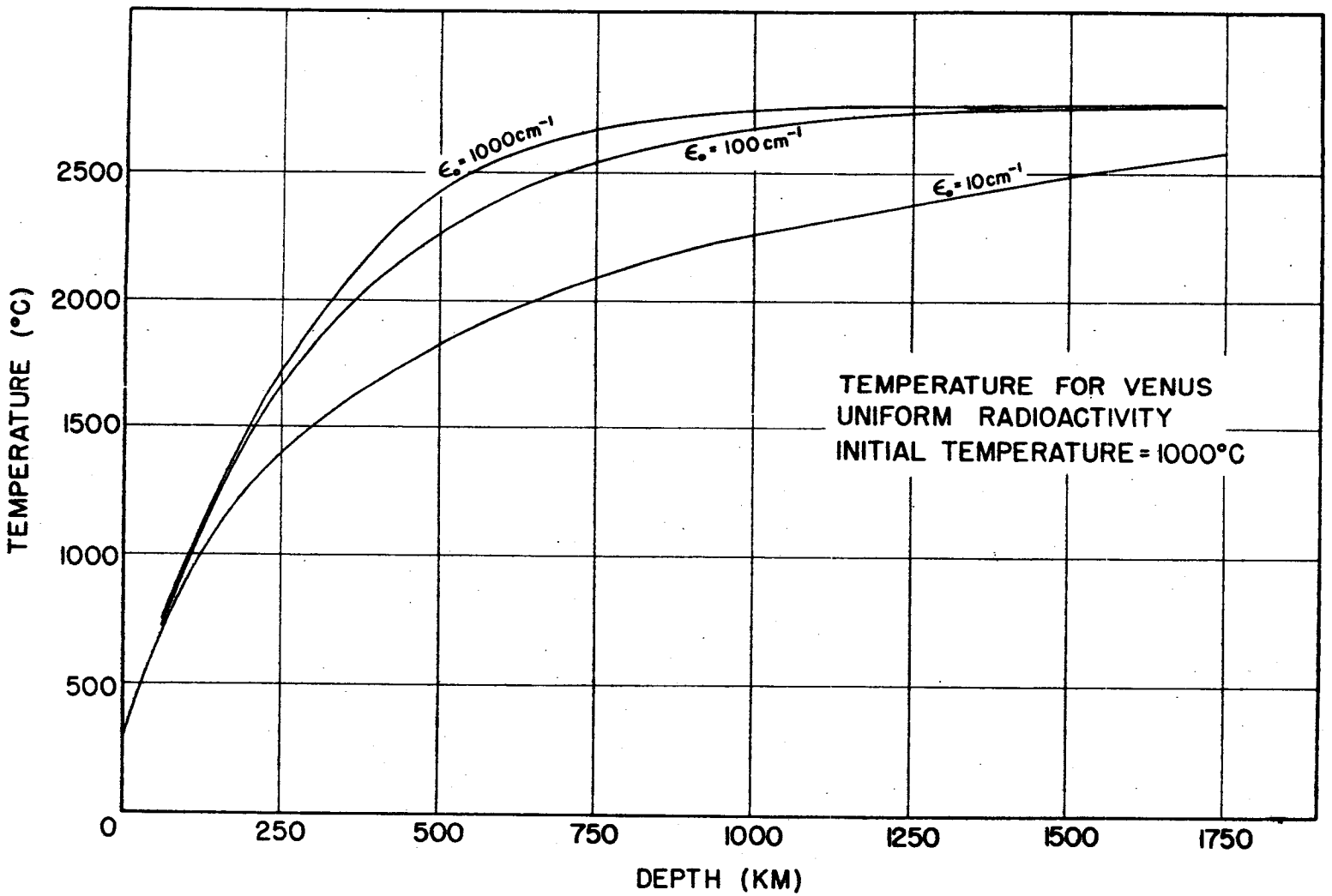


Figure 39: Calculated temperature distribution in Venus for an initial temperature of 1000°C. Chondritic radioactivity is assumed to be uniformly distributed. The detailed thermal parameters are listed in Table XVIII.

TABLE XVIII

Thermal Parameters for Models of Venus

Initial Temperature (uniform)	Venus I = 1000°C	Venus II = 500°C
Surface Temperature (constant in time)	300°C	
Radioactive Element Concentration (uniform)		
	$U = 1.1 \times 10^{-8}$ g g^{-1}	
	$Th = 4.4 \times 10^{-8}$ g g^{-1}	
	$K = 8.0 \times 10^{-4}$ g g^{-1}	
Time of integration	4.51×10^{9} years	
Lattice conductivity	2.5×10^{5} ergs cm^{-1} sec^{-1} $°C^{-1}$	
Density (uniform)	4.0 g cm^{-3}	
Heat capacity	1.3×10^{7} ergs g^{-1} $°C^{-1}$	

Final Surface Heat Flow (ergs cm^{-2} sec^{-1}) for	Venus I	Venus II
ϵ_o = 10 cm^{-1}	17.2	14.4
ϵ_o = 100 cm^{-1}	17.7	14.6
ϵ_o = 1000 cm^{-1}	16.7	14.0

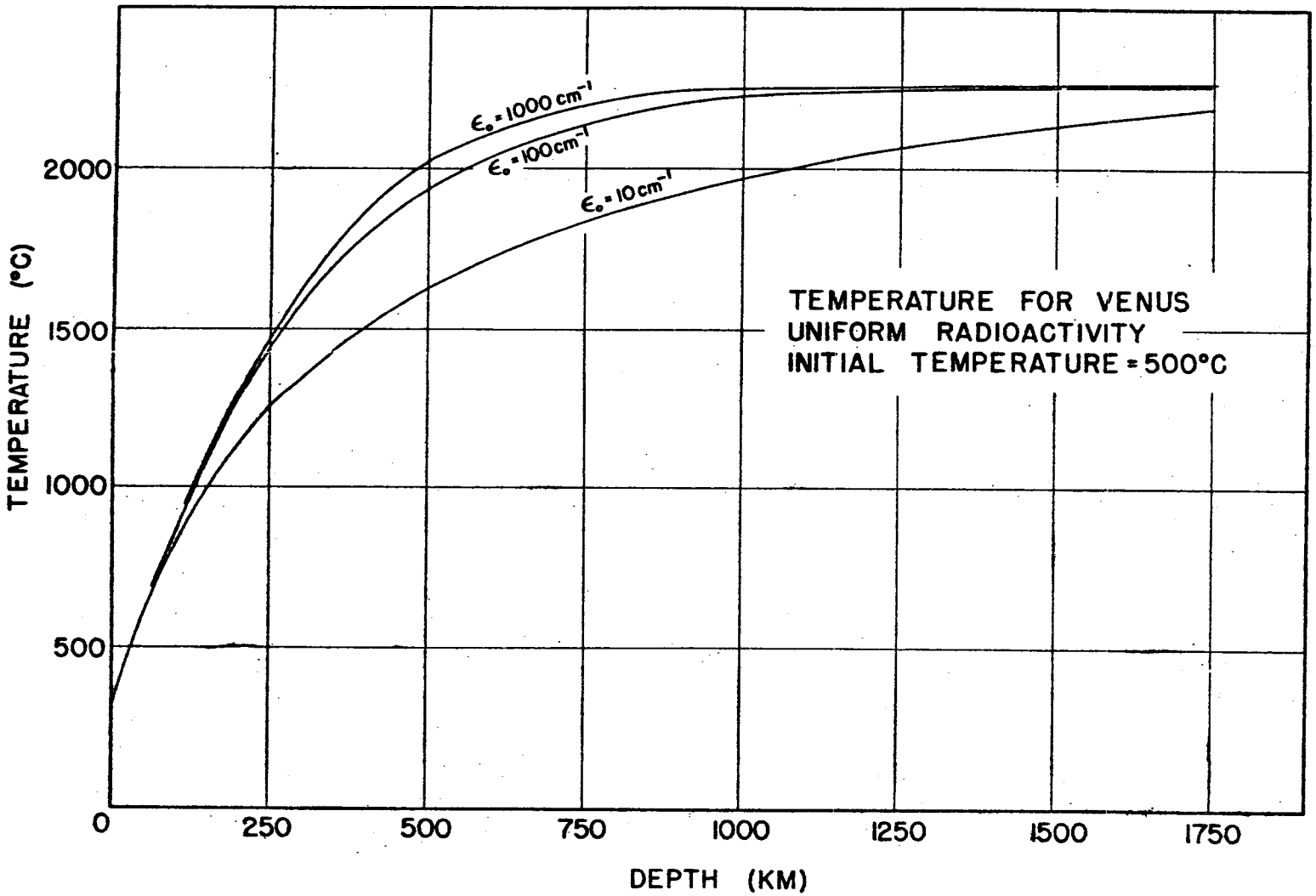


Figure 40: Calculated temperature distribution in Venus for an initial temperature of 500°C. Chondritic radioactivity is assumed to be uniformly distributed. The detailed thermal parameters are listed in Table XVIII.

Interior of Mercury

Inferences regarding the internal structure of Mercury are even more uncertain. The solar tide acting on Mercury is about 50 times more effective than the combined lunar and solar tides acting on the earth provided that the phase lag is the same. If Mercury had an initial angular momentum density proportional to its mass (see Figure 38), the time required to reduce its rate of rotation to the present value is long compared with the age of the solar system if the anelasticity of Mercury is equal to that of the earth. On the hypothesis of an initial angular momentum density equal to that of the other planets, we would require that either:

1) Mercury was initially molten and during this early time its period of rotation and revolution were equalized by solar tidal friction, or that

2) the present thermal state of Mercury is such as to allow a much greater dissipation of energy than in the earth.

The thermal conditions within Mercury are most uncertain. The mean density makes unlikely the hypothesis that the radioactive composition of Mercury is similar to that of chondrites. Indeed, it would seem that Mercury is largely metallic and if this is the case then a lower concentration of radioactive elements is indicated. Since the effect of pressure on the melting point is far less in a planet the size of Mercury than in larger bodies such as Venus and the earth, a lower concentration of radioactive elements, 1/3 to 2/3 that of the earth, would still make Mercury largely molten at present.

The Magnetic Field of Venus

The observations of Mariner II show that the dipole moment of Venus is at most 1/10 to 1/20 that of the earth. This observation can be explained either by supposing that:

1) Venus has no metallic core.

2) The origin of the field is precessional.

3) The field has a low dipole component but does contain higher order moments.

Hypothesis (1) is unlikely since the mass and dimensions are so similar to that of the earth. The thermal structures considered above would require that any iron present on the planet would segregate into the interior. The lack of a magnetic field may indicate the possibility of the precessional origin of the earth's field. Precessional torques acting on Venus are minute compared with those of the earth since the precessional torque is proportional to the square of the angular velocity. The third possibility cannot be ruled out. A field containing high-order moments in a non-rotating planet would not extend as far out into space as a dipole field and perhaps would have gone undetected by Mariner.

REFERENCES

Aldrich, L. T., and G. Wetherill, Geochronology by radioactive decay, Ann. Rev. Nucl. Sci., 8, 257-298, 1958.

Alexandrov, I., The lunar gravitational potential, in Advances in Astronomical Sciences, 5, Plenum Press, New York, 320-324, 1960.

Alsop, L. E., G. H. Sutton, and M. Ewing, The free oscillations of the earth observed on strain and pendulum seismographs, J. Geophys. Res., 66, 631-641, 1961.

Alterman, Z., H. Jarosch, and C. L. Pekeris, Oscillations of the earth, Proc. Roy. Soc., London, A., 252, 80-95, 1959.

Axelrod, D., Fossil flora suggest stable, not drifting continents, J. Geophys. Res., 68, 3257-3264, 1963.

Backus, G., and F. Gilbert, The rotational splitting of the free oscillations of the earth, Proc. Natl. Acad Sci., U.S., 47, 362-371, 1961.

Baldwin, J. E., Thermal radiation from the moon and heat flow through the lunar surface, Mon. Not. Roy. Astron. Soc., 122, 513-522, 1961.

Baldwin, R. B., The Measure of the Moon, University of Chicago Press, Chicago, 1962.

Barath, F. T., A. H. Barrett, J. Copeland, D. E. Jones, and A. E. Lilley, Microwave radiometers, Science, 139, 908-909, 1963.

Benioff, H., J. C. Harrison, L. LaCoste, W. H. Munk and L. B. Slichter, Searching for the earth's free oscillations, J. Geophys. Res., 64, 1334-1337, 1959.

Benioff, H., F. Press, and S. Smith, Excitation of the free oscillations of the earth by earthquakes, J. Geophys. Res., 66, 605-619, 1961.

Birch, F., Flow of heat in the Front Range, Colorado, Geol. Soc. Am. Bull., 61, 567-630, 1950.

Birch, F., Elasticity and constitution of the earth's interior, J. Geophys. Res., 57, 227-286, 1952.

Birch, F., The present state of geothermal investigation, Geophysics, 19, 645-659, 1954.

Birch, F., Differentiation of the mantle, Geol. Soc. Am. Bull., 69, 483-486, 1958.

Birch, F., Composition of the earth's mantle, Geophys. J., 4, 295-310, 1961.

Birch, F., and H. Clark, The thermal conductivity of rocks and its dependence upon temperature and composition, Am. J. Sci., 238, 529-558, 613-635, 1940.

Bolt, B., Spheroidal oscillations of the moon, Nature, 188, 1176-1177, 1961.

Bomford, G., Geodesy, Clarendon Press, Oxford, 1962.

Boyd, F. R., and J. L. England, Effect of pressure on the melting of diopside, $CaMgSi_2O_6$, and albite, $NaAlSi_3O_8$, in the range up to 50 kilobars, J. Geophys. Res., 68, 311-324, 1963.

Bullard, E. C., The magnetic field within the earth, Proc. Roy. Soc., London, A, 197, 433-453, 1949.

Bullard, E. C., The flow of heat through the floor of the Atlantic Ocean, Proc. Roy. Soc., London, A, 222, 408-429, 1954.

Bullard, E. C., A. E. Maxwell, and R. Revelle, Heat flow through the deep sea floor, in Advan. Geophys., 3, Academic Press, New York, 153-181, 1956.

Bullen, K. E., An Introduction to the Theory of Seismology, Cambridge University Press, London, 1953.

Camichel, H., Mesure de la dureé D de rotation de Saturn d'apre une Tache, Bull. Astronomique, 20, 141-144, 1954.

Chandler, S., On the variation of latitude, Astron. J., 11, 83, 1891.

Clark, S. P., Absorption spectra of some silicates in the visible and near infrared, Am. Mineralogist, 42, 732-742, 1957.

Clark, S. P., and E. R. Niblett, Terrestrial heat flow in the Swiss Alps, Mon. Not. Roy. Astron. Soc., Geophysical Supplement, 7, 176-195, 1956.

Clarke, F. W., and H. S. Washington, The composition of the earth's crust, U.S. Geological Survey, Professional Paper 127, 1924.

Clarke, V. C., Constants and related data used in trajectory calculations at the Jet Propulsion Laboratory, California Institute of Technology, Jet Propulsion Laboratory Technical Report, 32-273, 1962.

Cox, A., and R. R. Doell, A review of paleomagnetism, Geol. Soc. Am. Bull., 71, 645-768, 1960.

Dachille, F., and R. Roy, System Mg_2SiO_4 - Mg_2GeO_4 at 10,000, 60,000, and about 300,000 psi, Geol. Soc. Am. Bull., 67, 1682-1683, 1956.

Dachille, F., and R. Roy, Experimental study of the olivine-spinel inversion in Mg_2SiO_4, Geol. Soc. Am. Bull., 69, 1515, 1958.

Dachille, F., and R. Roy, High-pressure studies of the system Mg_2GeO_4-Mg_2SiO_4 with special reference to the olivine-spinel transition, Am. J. Sci., 258, 225-246, 1960.

Dolginov, Sh. Sh., E. G. Yeroshenko, L. N. Zhuzgof, and N. V. Pushkov, Investigation of the magnetic field of the moon, Geomagnetizm i Aeronomiya, 1, 21-29, 1961.

Dollfus, A., Un Micrometer, a double image permettant un grond dedublement, Compt. Rend., 235, 1479, 1952.

Dollfus, A., Mesure des dimensions du globe de la planéte Mars, Compt. Rend., 255, 2229-2231, 1962.

Elsasser, W. M., On the origin of the earth's magnetic field, Phys. Rev., 55, 489-498, 1939.

Gast, P. W., Limitations on the composition of the upper mantle, J. Geophys. Res., 65, 1287-1297, 1960.

Goldschmidt, D. M., Geochemistry, Clarendon Press, Oxford, 1954.

Goldstein, I. N., and R. L. Carpenter, Rotation of Venus: period estimated from radar measurements, Science, 139, 910-911, 1963.

Gutenberg, B., Physics of the Earth's Interior, Academic Press, New York, 1959.

Habibullin, S. T., On a unique solution for the f parameter of the physical liberation of the moon, Astron. J. Acad. of Sciences, USSR, 35, 669-680, 1958.

Hagihara, Y., Recommendations on the notation of the earth potential, Astron. J., 67, 108, 1962.

Hales, A. L., Convection current in the earth, Mon. Not. Roy. Astron. Soc., Geophysical Supplement, 3, 372-379, 1936.

Harrison, J. C., N. F. Ness, I. M. Longman, R. F. S. Forbes, E. A. Kraut, and L.B. Slichter, Earth tide observations made during the International Geophysical Year, J. Geophys. Res., 68, 1497-1516, 1963.

Heiskanen, W., and F. A. Vening-Meinesz, The Earth and It's Gravity Field, McGraw-Hill, New York, 1958.

Hess, H. H., History of ocean basins, in Petrologic Studies, A Volume in Honor of A. F. Buddington, Edited by A. E. J. Engle, H. L. James, B. L. Leonard, Geol. Soc. Am., 599-620, 1962.

Hoyle, F., and W. A. Fowler, On the abundances of uranium and thorium in solar systems material, in Isotopic and Cosmic Chemistry, North-Holland Publishing Company, Amsterdam, In press, 1963.

James, R., and Z. Kopal, The equilibrium figures of the earth and the major planets, Icarus, 1, 442-454, 1963.

Jeffreys, H., The structure of the earth down to the 20° discontinuity, Mon. Not. Roy. Astron. Soc., Geophysical Supplement, 3, 401-422, 1936.

Jeffreys, H., The density distribution in the inner planets, Mon. Not. Roy. Astron. Soc., Geophysical Supplement, 4, 62-71, 1937.

Jeffreys, H., The Earth, Cambridge University Press, 4th Edition, London, 1959.

Jeffreys, H., On the figure of the moon, Mon. Not. Roy. Astron. Soc., 122, 421-432, 1961.

Jeffreys, H., and R. Vicente, The theory of nutation and the variation of latitude, Mon. Not. Roy. Astron. Soc., 117, 142-161, 1957.

Kaula, W. M., Analysis of satellite observations for longitudinal variations of the gravitational field, in Space Research, 2, Proceedings of the Second International Space Science Symposium, Florence, North-Holland Publishing Company, Amsterdam, 360-372, 1961.

Kaula, W. M., Celestial geodesy, in Advan. Geophys., 9, Academic Press, 191-293, 1962.

Kaula, W. M., Tesseral harmonics of the gravitational field and geodetic datum shifts derived from camera observations of satellites, J. Geophys. Res., 68, 473-484, 1963a.

Kaula, W. M., Improved geodetic results from camera observations of satellites, J. Geophys. Res., In press, 1963b.

Kaula, W. M., A review of geodetic parameters, NASA Technical Note, D-1847, 1963c.

Kaula, W. M., Elastic models of the mantle corresponding to variations in the external gravity field, J. Geophys. Res., In press, 1963d.

Kaula, W. M., The investigation of the gravitational fields of the moon and planets with artificial satellites, in Advances in Space Science and Technology, Academic Press, New York, In press, 1963e.

King-Hele, D., The earth's gravitational potential deduced from the orbits of artificial satellites, Geophys. J., 4, 3-16, 1961.

King-Hele, D., The earth's gravitational potential deduced from the orbits of artificial satellites, Geophys. J., 6, 270-272, 1962.

Knopoff, L., and G. J. F. MacDonald, Attenuation of small amplitude stress waves in solids, Rev. Mod. Phys., 30, 1178-1192, 1958.

Knopoff, L., and G. J. F. MacDonald, An equation of state for the earth's core, Geophys. J., 3, 68-77, 1960.

Kopal, Z., Thermal history of the moon and terrestrial planets, 3. Convection and planetary interiors, Jet Propulsion Laboratory, California Institute of Technology, Technical Report No. 32-276, 1962.

Kopal, Z., Convection in planetary interiors, Icarus, 1, 391-400, 1963.

Kozai, Y., The gravitational field of the earth deduced from the motion of three satellites, Astron. J., 66, 8-10, 1961a.

Kozai, Y., Tesseral harmonics of the potential of the earth as derived from satellite motions, Astron. J., 66, 355-358, 1961b.

Kozai, Y., Numerical results from orbits, Smithsonian Institution, Astrophysical Observatory, Special Report No. 101, 1962a.

Kozai, Y., The potential of the earth derived from satellite motions, Proc. of the IUTAM Symposium, Dynamics of Satellites, Paris, Springer-Verlag, Berlin, In press, 1962b.

Kuenen, T. H., Marine Geology, J. Wiley, New York, 1950.

Kuhn, W., and A. Rittmann, Über den Zustand des Erdinnern und Seine Enstehung aus Einem Homogenen Urzustand, Geol. Rundschau, 32, 215-256, 1941.

Lee, W. H. K., Heat flow data analysis, Reviews of Geophysics, In press, 1963.

Lee, W. H. K., and G. J. F. MacDonald, Global variation of terrestrial heat flow, J. Geophys. Res., In press, 1963.

Levin, B. J., Thermal history of the moon, in The Moon, edited by Z. Kopal, Proceedings of Symposium No. 14 of the Internat'l Astronomical Union, Academic Press, 1962, pp. 157-167.

Love, A., A Treatise on the Mathematical Theory of Elasticity, Dover Publications, New York, 1927.

Lubimova, H. A., Thermal history of the earth with consideration of the variable thermal conductivity of the mantle, Geophys. J., 1, 115-134, 1958.

MacDonald, G. J. F., Chondrites and the chemical composition of the earth, in Researches in Geochemistry, Edited by P. H. Abelson, J. Wiley, New York, 476-494, 1959a.

MacDonald, G. J. F., Calculations on the thermal history of the earth, J. Geophys. Res., 64, 1967-2000, 1959b.

MacDonald, G. J. F., Stress history of the moon, Planetary Space Sci., 2, 249-255, 1960.

MacDonald, G. J. F., Interior of the moon, Science, 133, 1045-1050, 1961.

MacDonald, G. J. F., The internal constitution of the inner planets, J. Geophys. Res., 67, 2945-2974, 1962.

MacDonald, G. J. F., The deep structure of continents and oceans, Reviews of Geophysics, In press, 1963.

MacDonald, G. J. F., and L. Knopoff, Chemical composition of the outer core, Geophys. J., 1, 284-297, 1958.

MacDonald, G. J. F., and N. F. Ness, A study of the free oscillations of the earth, J. Geophys. Res., 66, 1865-1911, 1961.

Malkus, W. V. R., Precessional torques as the cause of geomagnetism, J. Geophys. Res., 68, 2871-2886, 1963.

Mayer, C., Radioemission from the moon and planets, in Planets and Satellites, Chicago University Press, Chicago, 442-472, 1961.

Molodenskii, M. S., The theory of nutation and daily earth tides, Bull. (Izvestiya) Acad. Sci., USSR, Geophysics Series, 1961.

Munk, W. H., and E. S. Hassan, Atmospheric excitation of the earth's wobble, Geophys. J., 4, 339-358, 1961.

Munk, W. H., and G. J. F. MacDonald, The Rotation of the Earth, Cambridge University Press, London, 1960a.

Munk, W. H., and G. J. F. MacDonald, Continentality and the gravitational field of the earth, J. Geophys. Res., 65, 2169-2172, 1960b.

Ness, N. F., J. C. Harrison, and L. B. Slichter, Observations of the free oscillations of the earth, J. Geophys. Res., 66, 621-629, 1961.

Newton, R., Ellipticity of the equator deduced from the motion of Transit 4A, J. Geophys. Res., 67, 415-416, 1962.

O'Keefe, J., and C. D. Batchlor, Perturbations of a close satellite by the equatorial ellipticity of the earth, Astron. J., 62, 183-185, 1957.

O'Keefe, J. and W. C. Cameron, Evidence from the moon's surface features for the production of lunar granites, Icarus, 1, 271-285, 1962.

O'Keefe, J., A. Eckels, and K. Squires, The gravitational field of the earth, Astron. J., 64, 245-253, 1959.

Pariisky, N. N., Studies of tidal variations of gravity in central Asia by the Earth Physics Institute of the USSR Academy of Sciences, Quatrieme Symposium Intern. sur les Marees Terrestres, Bruxelles, 5-10 Jiun 1961, Communications de l'Observ. R. de Belgique, No. 188, S. G. No. 58, 1961.

Pariisky, N. N., Earth tides and the internal structure of the earth, Bull. (Izvestiya) Acad. Sci., USSR, Geophysics Series, 193-215, 1963.

Pekeris, C. L., Thermal convection in the interior of the earth, Mon. Not. Roy. Astron. Soc., Geophysical Supplement, 3, 340-367, 1935.

Pekeris, C. L., Z. Alterman, and H. Jarosch, Rotational multiples in the spectrum of the earth, Phys. Rev., 122, 1692-1700, 1961a.

Pekeris, C. L., Z. Alterman, and H. Jarosch, Comparison of theoretical with observed values of the periods of free oscillations of the earth, Proc. Natl Acad. Sci., U.S., 47, 91-98, 1961b.

Pettengill, G. H., H. W. Briscoe, J. V. Evans, E. Gehrels, G. M. Hyde, L. G. Kraft, R. Price and W. B. Smith, A Radar Investigation of Venus, Astron. J., 67, 181-190, 1962.

Rabe, W., Untersuchungen über die Durchmesser der Grossen Planeten, Astron. Nachr., 234, 153, 1926.

Ramsey, W. H., On the constitution of the terrestrial planets, Mon. Not. Roy. Astron. Soc., 108, 406-413, 1948.

Ramsey, W. H., On the nature of the earth's core, Mon. Not. Roy. Astron. Soc., Geophysical Supplement, 5, 409-426, 1949.

Ramsey, W. H., The planets and the white dwarfs, Mon. Not. Roy. Astron. Soc., 110, 444-454, 1950.

Rankama, K., and T. G. Sahama, Geochemistry, University of Chicago Press, Chicago, 1950.

Ringwood, A. E., The system Mg_2SiO_4 - Mg_2GeO_4, Am. J. Sci., 254, 707-711, 1956.

Ringwood, A. E., The system Mg_2SiO_4 - Mg_2GeO_4, Geol. Soc. Am. Bull., 69, 129, 1958a.

Ringwood, A. E., The system Mg_2SiO_4 - Mg_2GeO_4 , Geochim. Cosmochim. Acta, 15, 18-29, 1958b.

Ringwood, A. E., Olivine-spinel transition in fayalite, Geol. Soc. Am. Bull., 69, 129-130, 1958c.

Ringwood, A. E., and N. Seabrook, High-pressure transition of $MgGeO_3$ from pyroxene to corundum structure, J. Geophys. Res., 67, 1690-1691, 1962.

Ross, C. S., M. D. Foster, and A. T. Meyers, Origin of dunite and olivine rich inclusions in basaltic rocks, Am. Mineralogist, 39, 693-757, 1954.

Runcorn, S. K., Paleomagnetic evidence for continental drift and its geophysical cause, in Continental Drift, Edited by S. K. Runcorn, Academic Press, New York, 1-40, 1962a.

Runcorn, S. K., Towards a theory of continental drift, Nature, 193, 311-314, 1962b.

Runcorn, S. K., Physics of the moon, a geophysical discussion, Quart. J., Roy. Astron. Soc., 4, 85, 1963.

Russell, H. N., R. S. Dugan, and J. Q. Stewart, Astronomy I - The Solar System, Ginn and Company, Boston, 1945.

Smith, E. J., L. Davis, P. J. Coleman, and C. P. Sonett, Magnetic field of Venus, Science, 139, 909-910, 1963.

Takeuchi, H., M. Saito, and N. Kobayashi, Free oscillations of the moon, J. Geophys. Res., 66, 3895-3897, 1961.

Tilton, J. R., and G. L. Davis, Geochronology, in Researches in Geochemistry, Edited by P. H. Abelson, J. Wiley, New York, 190-216, 1959.

Tilton, J. R., and G. W. Reed, Radioactive heat production in eclogite and some ultramafic rocks, in Earth Science and Meteoritics, North-Holland Publishing Company, Amsterdam, 31-44, 1963.

Torreson, O. W., T. Murphy, and J. W. Graham, Magnetic polarization of sedimentary rocks and the earth's magnetic history, J. Geophys. Res., 54, 111-129, 1949.

Trumpler, R. J., Observations of Mars at the opposition of 1924, Lick Observatory Bulletin, 13, 19-45, 1926.

Urey, H. C., The Planets, Their Origin and Development, Yale University Press, New Haven, 1952.

Urey, H. C., Boundary conditions for theories of the origin of the solar system, in Phys. Chem. Earth, 2, Pergamon Press, New York, 46-47, 1957.

Urey, H. C., Lines of evidence in regard to the composition of the moon, in Space Research, North-Holland Publishing Company, Amsterdam, 1960.

Urey, H. C., Origin and history of the moon, in Physics and Astronomy of the Moon, Edited by Z. Kopal, Academic Press, New York, 481-524, 1962.

Vening-Meinesz, F. A., Gravity expeditions at sea, Delftsch Uitgevers Maatschappy, Delft, 4, 1923-1938, 1948.

Vening-Meinesz, F. A., Thermal convection in the earth's mantle, in Continental Drift, Edited by S. K. Runcorn, Academic Press, New York, 145-176, 1962.

Verhoogen, J., The heat balance of the earth's core, Geophys. J., 4, 276-281, 1961.

Vestine, E. H., L. Laporte, C. Cooper, I. Lang, and W. C. Hendrix, Description of the earth's magnetic field and its secular changes, Carnegie Institute of Washington, Publication 578, 1947.

Wilson, J. T., Some consequences of the expansion of the earth, Nature, 185, 880-882, 1960.

Woolard, E. W., The secular perturbations of the satellites of Mars, Astron. J., 51, 33-36, 1944.

Worzel, J. L., and G. L. Shurbet, Gravity interpretations from standard oceanic and continental cross sections, in Crust of the Earth, Edited by A. Poldervaart, Geol. Soc. Am., Spec. Papers 62, 87-100, 1954.

Yakovkin, A. A., The free librations of the moon, Transactions IAU, 8, 231-232, 1952.

THE ESCAPE OF HELIUM FROM THE EARTH'S ATMOSPHERE

by

Gordon J. F. MacDonald

Institute of Geophysics
and Planetary Physics
University of California
Los Angeles, California

THE ESCAPE OF HELIUM FROM THE EARTH'S ATMOSPHERE

TABLE OF CONTENTS

ABSTRACT 1

I. INTRODUCTION 3

II. MODELS OF THE UPPER ATMOSPHERE 10

Data Obtained from Satellite Drag Studies 10

Construction of Numerical Models 16

Accession of Heat 22

Influence of Horizontal Temperature Gradients 28

Ultraviolet Heat Source 30

Radiative Loss of Heat 35

Numerical Models of the Upper Atmosphere 36

Boundary conditions at 120 km 37

Required heat sources 39

Numerical models 41

Solar-cycle variations 43

Summary of Atmospheric Constitution 44

III. FLUX OF CONSTITUENTS INTO THE ATMOSPHERE 48

Introduction 48

The Flux of Helium 4 49

The Flux of Helium 3 53

The Accretion of Neon 54

Summary of Flux of Gases into Atmosphere 56

IV. THERMAL ESCAPE OF GASES 58

Introduction 58

Classical Theory of Escape 59

Theory of Diffusion of a Minor Component 63

Escape of Helium from the Earth's Atmosphere 67

V. EFFECT OF THE MAGNETIC FIELD 70

Introduction 70

Direct Effect of the Field 72

Thermal Processes Associated with the Magnetic Field 73

VI. SUMMARY AND CONCLUSIONS 79

REFERENCES 82

LIST OF TABLES

TABLE I	-	24
TABLE II	- Flux of Solar Radiation at Short Wavelengths	32
TABLE III	- Concentration of Constituents at 120 Km	40
TABLE IV	- Maximum, Mean, and Minimum Daily Temperatures	45
TABLE V	- Concentration of Rare Gases in the Earth's Atmosphere	50
TABLE VI	- Flux of Gases into the Atmosphere	57

LIST OF FIGURES

	Figure Heading	After Page
Figure 1	The power spectrum of the drag on Echo 1 and of the 20-cm solar flux.	11
Figure 2	The coherence between the 20-cm solar flux and drag on Echo 1.	11
Figure 3	The power spectrum of the drag on Explorer 9 and the 10.7-cm solar flux.	11
Figure 4	Coherence between drag on Explorer 9 and 10.7-cm solar flux.	11
Figure 5	The number density of neutral molecules at 0 and 12 hours local time as a function of altitude.	41
Figure 6	The number density of neutral particles in the high atmosphere at 6 and 18 hours local time.	41
Figure 7	The relative abundance of neutral particles as a function of altitude at 0 hours local time.	41
Figure 8	The relative abundance of neutral particles at 6 hours local time.	41
Figure 9	A relative abundance of neutral particles at 12 hours local time.	41
Figure 10	A relative abundance of neutral particles at 18 hours local time.	41
Figure 11	The temperature at 0 and 12 hours local time.	43
Figure 12	Temperatures at 6 and 18 hours local time.	43

THE ESCAPE OF HELIUM FROM THE EARTH'S ATMOSPHERE

Gordon J. F. MacDonald
Institute of Geophysics
and Planetary Physics
University of California
Los Angeles, California

ABSTRACT

The rate of influx of helium 4 from the earth's crust into the atmosphere is about 2×10^6 atoms cm^{-2} sec^{-1}. The average rate of thermal escape averaged over the last solar cycle is 6×10^4 atoms cm^{-2} sec^{-1}. The present abundance of helium in the atmosphere would accumulate in about two million years. Either some nonthermal mechanism permits at present a greater rate of escape of helium or the rate of escape has varied in time. This paper is concerned with an examination of possible changes in the escape rate due to variations in the earth's magnetic field. Problems associated with the heat balance in the upper atmosphere are reviewed. Statistical analysis of drag on satellites and solar decimeter flux establishes that the 27-day variation in atmospheric density results from fluctuations in the intensity of the solar wind. It is considered likely that a heat source over and beyond that provided by ultraviolet radiation is required to account for both the day-night variations in upper atmosphere densities and the 27-day variation. Various mechanisms by which energy density in the solar wind can be transferred into the upper atmosphere are considered. It is suggested that, if the additional

heat source is indeed due to the solar plasma, historical fluctuations in the earth's magnetic field could result in a greater escape rate of helium. The establishment and identification of the additional heat source is necessary for a full understanding of the escape of helium. The terrestrial economy of helium 3 is reviewed. The contribution by solar cosmic rays to the abundance of helium 3 remains a major problem. The rate of influx of helium 3, about 10 atoms cm^{-2} sec^{-1} is greater than the escape rate averaged over the last solar cycle, about 4 atoms cm^{-2} sec^{-1}, if solar cosmic rays contain about 10 percent helium 3.

I. INTRODUCTION

Bates and McDowell (1957, 1959) first suggested that helium (He^4) in the atmosphere is not in equilibrium, but that instead the current rate of influx exceeds the escape flux. Better estimates of the generation rate have been obtained in recent years, and detailed investigations of the upper atmosphere provide improved limits on the rate of thermal escape. These new data clearly show that at present the rate at which helium 4 is supplied to the atmosphere greatly exceeds the rate of thermal escape. Bates and McDowell (1957) tentatively suggested that at some time in the past the escape rate was greater and pointed to the possibility of changes in the solar flux. Bates and Patterson (1962) made an alternative suggestion that helium is now in equilibrium but that the escape is nonthermal, involving the dissociative recombination of HeO^+.

Nicolet (1957) demonstrated that thermal escape could not account for the observed abundances of the two isotopes of helium, He^4 and He^3. If the temperature were high enough to account for the loss of He^4, then He^3 would be removed at a rate far in excess of its production. If He^3 escapes thermally and is in equilibrium, then He^4 remains trapped. The implication of Nicolet's (1957) study is that either steady equilibrium does not hold or there is some nonthermal process removing helium from the atmosphere.

The escape of helium is a complex problem, and despite advances there are many more problems than there are answers. The complexity requires a review of many aspects of geophysics. The escape problem forms a framework in which a number of fascinating aeronomical problems can be examined.

The rate of thermal escape of helium is a delicate function of the temperature at the escape level. This level, defined as the height at which the mean free path of helium becomes large compared to its scale height, is at 500-600 km. Small changes in the input of energy to the tenuous gas at high altitudes will make a large difference in the effective temperature. It is known that the ultraviolet flux, which largely determines the temperatures at these heights, does undergo changes with the solar cycle, and there may be longer-period fluctuations.

It is also known that the earth's magnetic field undergoes a variety of fluctuations. Indeed, of all internal fields, it is the most changeable. The nondipole part of the field undergoes long-period changes with an average value of 5×10^{-4} gauss year^{-1} and maximum rates of change of 1.5×10^{-3} gauss year^{-1}. The dipole field has changed over the past century at an average rate of 3.5×10^{-5} gauss year^{-1}, which implies that if the field changed linearly with time then ten thousand years ago the field was twice its current value. Indeed, studies of remnant magnetism in Roman ceramic objects suggest a field with twice

the present intensity only two thousand years ago (Hide and Roberts, 1961). These relatively rapid changes, combined with the controversial paleomagnetic evidence of a changing field on a geologic time scale, raise the problem of the effect of the magnetic field on the long-term stability of the upper atmosphere.

The long-term stability of a planetary atmosphere will certainly be influenced by variations in solar ultraviolet flux. The magnetic field may also influence the composition of the atmosphere at high altitudes. The effects of the changing solar flux can be treated in a relatively straightforward fashion. The influence of the magnetic field is complex and requires detailed examination. It is clear that on the earth an escaping neutral atom passes through the upper atmosphere much too fast for it to be ionized and trapped by the field. On the other hand, the magnetic field shields the upper atmosphere from the low-energy solar plasma and prevents this potential heat source from reaching it. The magnetic field also interacts with the solar plasma, and this interaction gives rise to hydromagnetic and electromagnetic waves that may in part determine the thermal conditions near the escape layer. In addition, the interaction may produce energetic electrons which enter the magnetosphere to deposit their energy in high-latitude regions. The shielding by the field also influences the accretion rate; this rate is important in a consideration of the economy of helium 3 and tritium.

The atmospheres of Venus and of the earth present two contrasting examples of development. At present, it is widely believed that the earth's atmosphere has evolved gradually through geologic time (Rubey, 1951, 1955; Holland, 1962). The continued exhalation of the gases from the earth's interior results from gradual heating of parts of the earth's interior. The rate of exhalation is controlled primarily by the internal thermal processes, which are as yet poorly understood. The internal structure of Venus is similar to that of the earth (MacDonald, 1962). The two planets have approximately the same linear dimensions and approximately the same mass. The bulk chemical composition of the two planets thus cannot differ to any great extent. It is also likely that the radioactive composition is similar. Thus, it is probable that Venus and the earth have had similar, though not necessarily identical, thermal histories. If this is so, there is no reason to expect that the gases, and particularly the rate at which the gases have been supplied to the outer surface of Venus, should differ from the rates that have held on the earth. The fact that the atmospheres apparently differ presents a number of fascinating problems.

One marked difference between the earth and Venus is that the magnetic field of Venus is weak compared with the earth and may be non-existent (Smith, et. al., 1963). Is this striking first-order difference between the two planets of possible importance in the divergent development of the planetary atmosphere?

The rapid advances in the understanding of the upper atmosphere make profitable a review of the current state of the theory of the escape of the atmosphere. Particular attention is given to the possibility of changes of escape rates in the past. In Section II we review the constitution of the upper atmosphere. A detailed study of atmospheric models is essential to the escape problem, since only through the model is an estimate of the temperature possible. Data on the drag of satellites provide reliable estimates of atmospheric density from 200 km up to at least 1000 km. These densities can then be used as controls on calculated models of the atmosphere. Construction of a model atmosphere would be impossible without detailed knowledge of the nature of the solar flux, since it is this flux that provides a major heat source. Rocket measurements of the extreme ultraviolet taken during the past five years have greatly clarified the position. The compositional models of the atmosphere must remain most uncertain because of the lack of knowledge of the most critical region within the atmosphere, 100-200 km. The convective equilibrium changes into a diffusive equilibrium with large temperature gradients. This region has only been perfunctorily explored by rockets. Conditions assumed here fix in a large part the density, composition, and temperature of the atmosphere at greater altitudes. Despite the uncertainties, it appears from model studies that ultraviolet radiation alone cannot account for the observed temporal variations in the density of the high atmosphere. Further evidence for a heat source in addition to solar radiation is found in the detailed comparison of the variation of solar radio flux and drag.

Section III reviews new data on the flux of the critical constituents into the atmsophere. It appears probable that particles emitted by solar flares add substantial fractions of helium 3 and tritium to the atmosphere. New studies on the chemical composition of natural gases permit a better estimate of the rate at which important constituents enter the atmosphere through its lower boundary.

Section IV describes the current state of the theory of escape from a planetary atmosphere. Considerable advances have been made since the early work of Jeans. Furthermore, it is clear that an important factor in determining the long-term stability of minor constituents is the rate at which a minor component diffuses through the major components. The diffusion rate is complicated by the phenomenon of thermal diffusion. A comparison of the present escape of helium 4 with production rates confirms Bates and McDowell's and Nicolet's conclusions about the helium balance in the atmosphere.

In Section V the effects of the magnetic field are considered. A strong field hinders the accretion and can influence the thermal balance. The additional heat source referred to above may be due to hydromagnetic and electromagnetic radiation resulting from the interaction of the solar wind with the magnetic field. If this is so, an increase in the past of the field would have resulted in a greater interaction cross section of the earth with the solar wind and a higher upper atmosphere temperature.

Alternatively, if no heat source over that provided by the ultraviolet radiation is required, then a weakening of the field could have provided a higher temperature in the past. A weak field would allow the penetration of the plasma and the deposition of thermal energy in the upper atmosphere. If the heating of the upper atmosphere is due entirely to ultraviolet radiation, the disappearance of the field would result in a greater rate of helium escape. Alternatively, if an additional heat source involving the interaction of the field with the plasma is important, then a strengthening of the field yields a higher escape rate. It is argued that the strengthening of the field is more likely, and that over at least 10 percent of geologic time the field was at least twice its present strength.

Section VI summarizes the many problems associated with the helium balance.

Because of the complexity of the subject, a guide to the organization of the review may be helpful. The helium problem is considered explicitly in Section III and the last part of Section IV. Sections II and IV are in the nature of a more general review of atmospheric constitution and can be omitted by the reader interested primarily in helium escape. Section V is a more speculative treatment of both heat sources in the upper atmosphere and of possible changes in these heat sources during geologic time. Summaries of the longer sections (II and III) are provided.

II. MODELS OF THE UPPER ATMOSPHERE

Data Obtained from Satellite Drag Studies

As a satellite passes through the upper atmosphere it interacts with the neutral gas. In colliding with the neutral particles the satellite transfers part of its momentum. The rate of transfer of momentum is greatest where the density of neutrals is high. Since the density in the atmosphere is approximately exponential, most of the momentum transfer takes place when the satellite dips farthest into the atmosphere. Observation of the secular change in the orbital elements permits a determination of atmospheric density in the region near perigee. Many studies have been carried out relating the atmospheric properties to changes in the orbits of satellites; these have been reviewed by, for example, Harris and Priester (1962a), Nicolet (1961a), Paetzold (1961), and King-Hele and Walker (1961).

The determination of the absolute value of density at a given height is complicated by the dependence of the drag parameter on the mass of a satellite, the effective cross section, and the satellite's drag coefficient. Since all these parameters vary from satellite to satellite, the absolute values of density derived from satellite observation remain in doubt. For high altitudes above 1000 to 1500 km, there is a further possible complication that interaction with ionized gas and the magnetic field may contribute to the drag. The separation of the electromagnetic contribution from the drag of the neutral molecules is difficult. A first attempt at confirming the density obtained from drag observations by direct measurement was carried out by Sharp et al. (1962) at 500 - 600 km. They used a ribbon microphone and converted

the observed ram pressure to densities. The values obtained agree well with those derived from orbits of satellites.

Though there is still some uncertainty regarding the absolute densities at high altitudes, their time variation is now well established. The spectrum of variations, though complicated, includes a strong line at 24 hours corresponding to a diurnal density variation. Priester et al. (1960) showed that at 200 km the variation represented a few percent of the mean drag. At 650 km the variation reaches almost a factor of 10 (Jacchia, 1960; Priester and Martin, 1960; King-Hele and Walker, 1960). Sharp et al. obtained a day-night variation of approximately a factor of 4 at 550 km.

High-frequency variations in the solar atmosphere are evidenced by the behavior of satellites during solar events. The most dramatic of these was the series of flares of November 1960 in which there were rapid changes in the density of the upper atmosphere (Jacchia, 1961; Groves, 1961; Bryant, 1961).

The analysis of drag of various satellites also shows a peak in the spectrum centered around 27 days and correlated with the variation in radio emission from the sun in the wavelength range 3-30 cm (Jacchia, 1958, 1959, 1960; Priester, 1961). The striking correlation between the radio flux and the variation in drag is illustrated in Figures 1-4. Zadunaisky et al. (1961) carried out a detailed analysis of the orbit of Echo I. The power spectrum of the Echo drag is shown in Figure 1

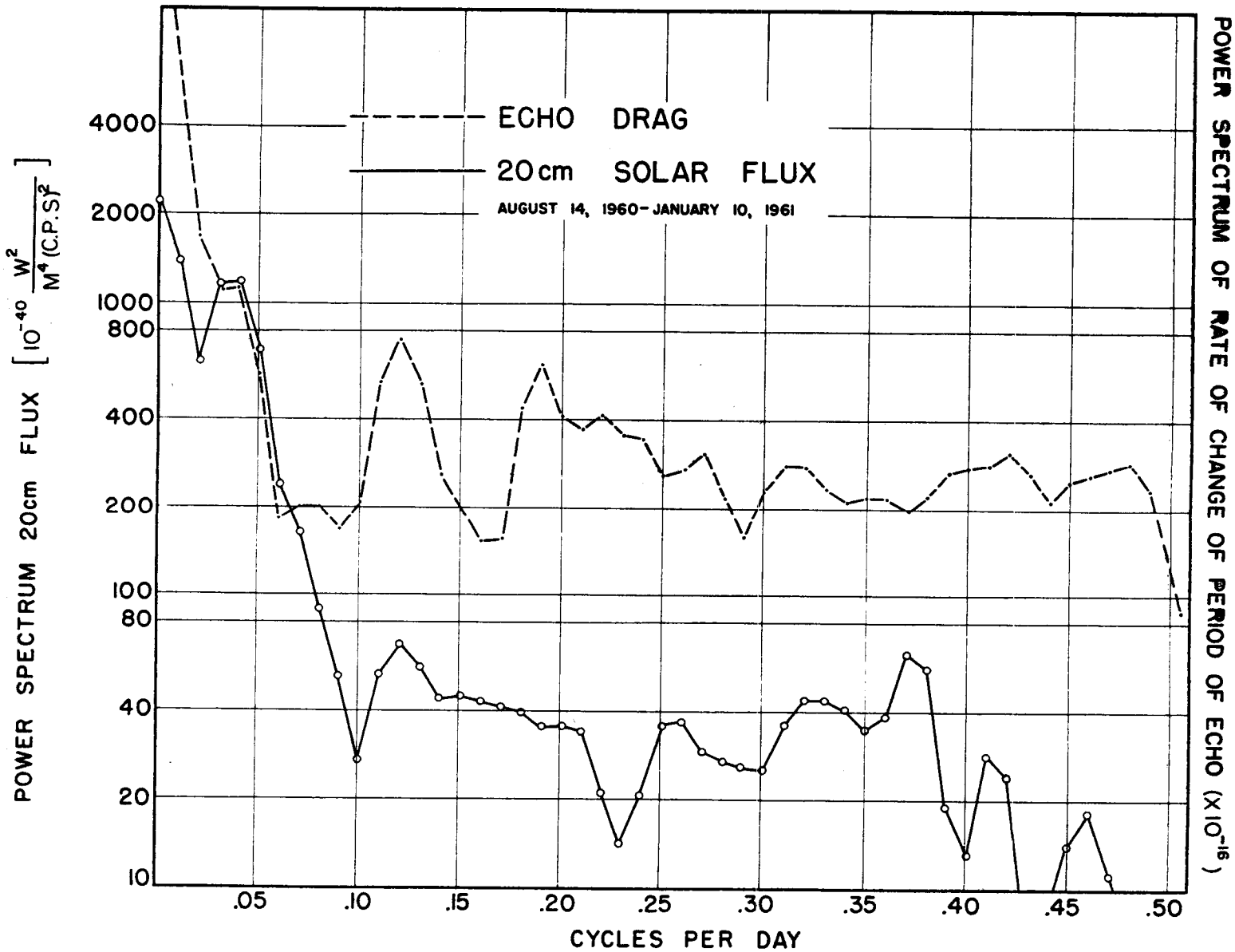


Figure 1: The power spectrum of the drag on Echo 1 and of the 20-cm solar flux. Power spectrums cover the interval August 14, 1960, to January 1961. Both spectrums climb at low frequencies, but the low-frequency components have purposely not been filtered out so as not to exaggerate the 27-day peak.

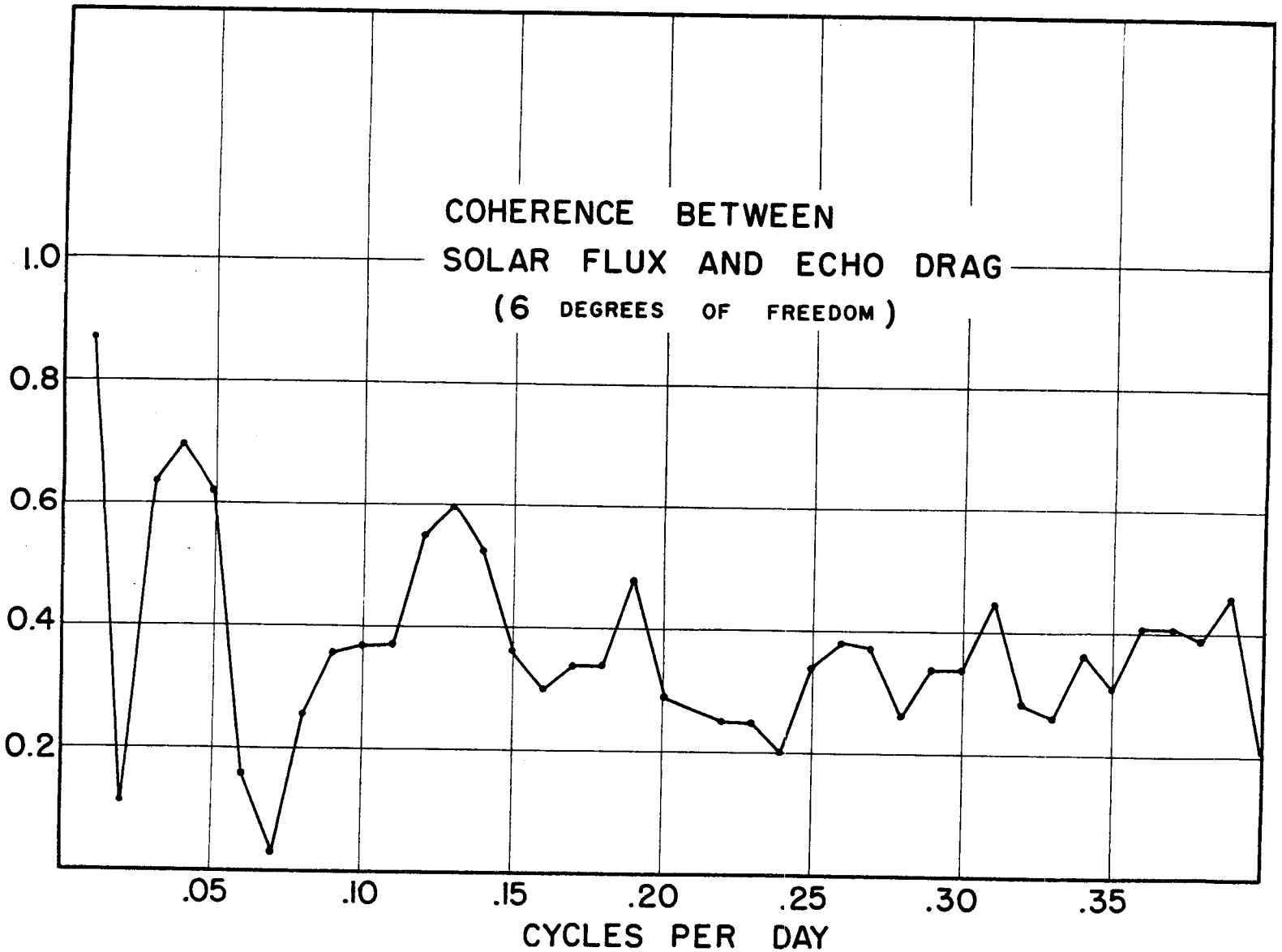


Figure 2: The coherence between the 20-cm solar flux and drag on Echo 1.

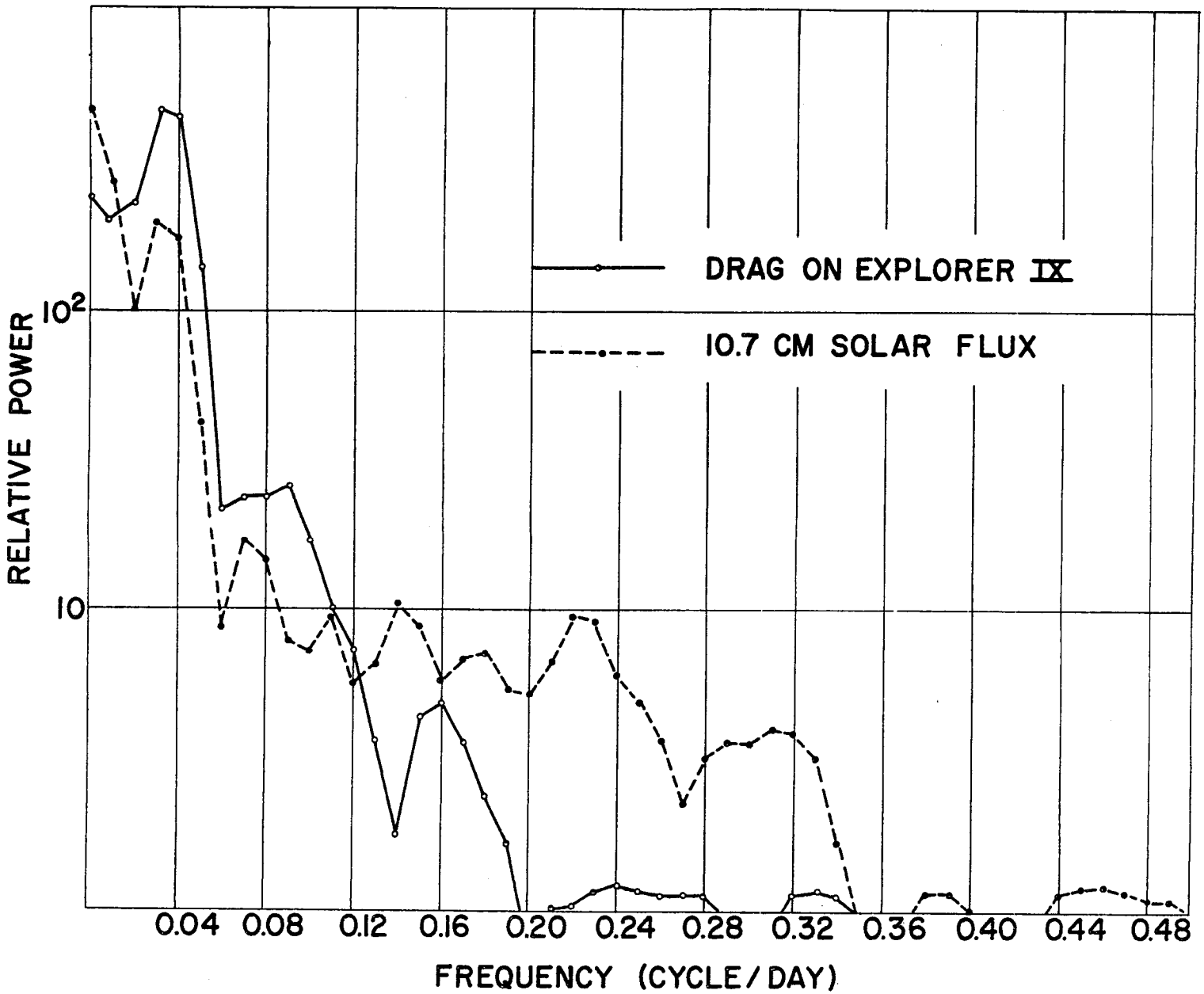


Figure 3: The power spectrum of the drag on Explorer 9 and the 10.7-cm solar flux.

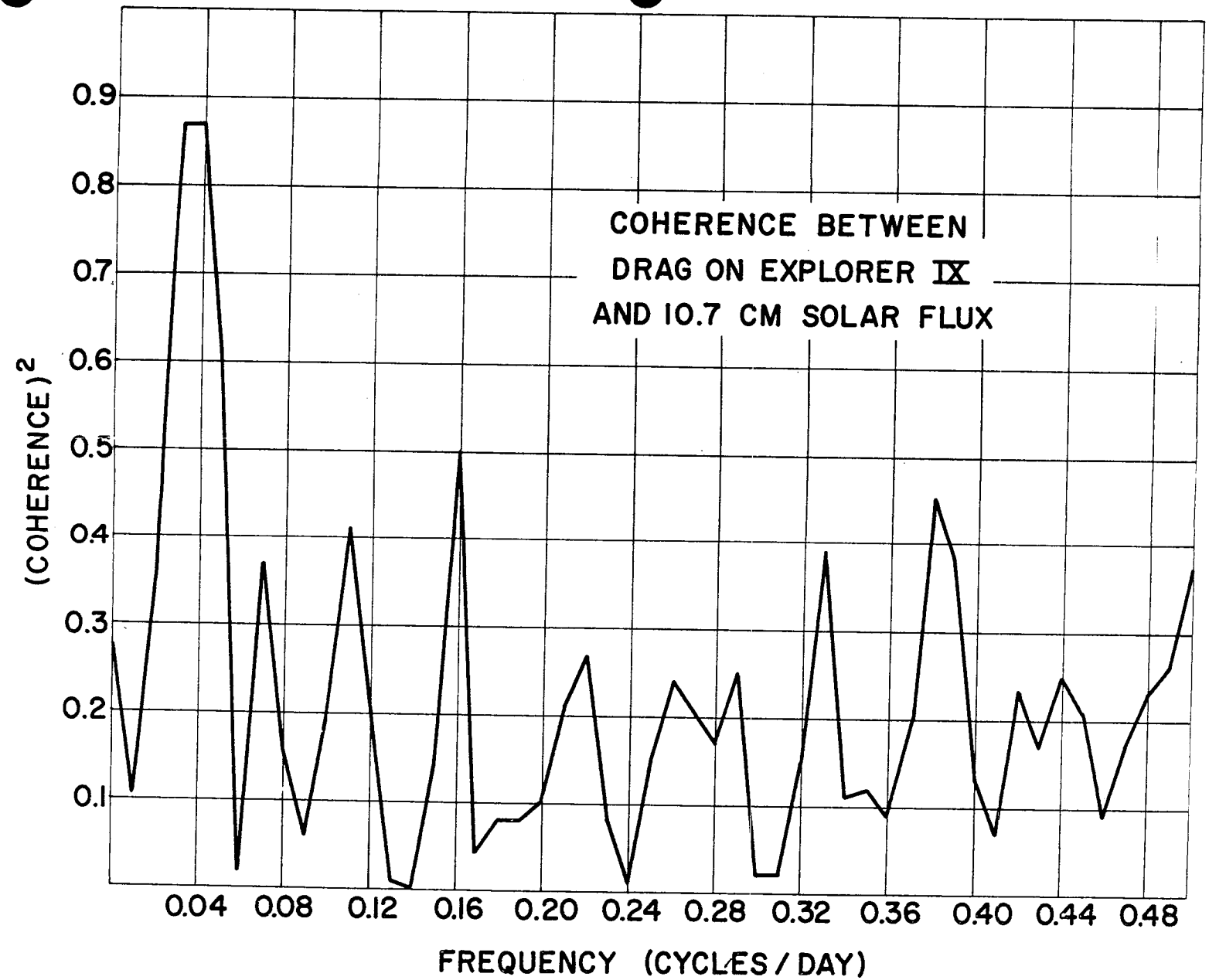


Figure 4: Coherence between drag on Explorer 9 and 10.7-cm solar flux. The only significant coherence is at the frequency corresponding to the 27-day period.

for the period August 13, 1960, to January 11, 1961. There is a definite peak at 0.037 cycle per day (27-day period) though the spectrum climbs rapidly at the low frequencies. The very-low-frequency variations have purposely not been filtered out, as filtering would exaggerate the 27-day peak. In addition, there is a peak at 0.12 cycle per day corresponding to a period of about 8 days. The power spectrum for the time variations of the 20-centimeter solar flux is also shown in Figure 1. Both the 27-day and the 8-day peaks appear in the spectrum. Figure 2 shows the coherence between the 20-centimeter flux and the drag. The coherence is 0.7 at the 27-day peak and 0.6 for the 8-day peak.

Explorer 9 was a 12-foot balloon satellite launched specifically to investigate the values of atmospheric drag. The orbital elements for the interval February 17 to October 2, 1961, have been analyzed in detail by Jacchia and Slowey (1962). The perigee height varied between 640 and 753 km over the 7.5-month interval that they studied. The power spectrums of the drag after removal of radiation pressure effects is shown in Figure 3. The drag on Explorer 9 is highly correlated ($R^2 = 0.87$) with the variation of the 10.7-centimeter radio flux (see Figures 3 and 4) at the 27-day period. There is some evidence for the 8-day period noted in the Echo 1 spectrums, but the peak is much weaker.

Though the solar decimeter flux is well correlated with changes in drag, the solar decimeter radiation cannot be the physical cause of the fluctuations, but is merely associated with the cause. The decimeter radiation is generally attributed to thermal emission from condensations in the solar corona. The variations in atmospheric density are thought to be primarily the result of variations in the extreme ultraviolet radiation. The ultraviolet radiation may arise from the coronal condensations, since emission is probably due to cascades following recombination on excited levels (Harris and Priester, 1962a).

If the changes in drag are due to ultraviolet radiation from coronal condensations, the time interval between an event on the sun and a variation in drag should equal the light transit time plus the response time of the atmosphere. We will later show that the response time is at most on the order of a few hours because of the short thermal time constant of the atmosphere. Thus the variation of drag should lag behind the fluctuation in decimeter flux by about two hours if the heating is due to ultraviolet radiation.

A direct determination of the phase relations would require the identification of a solar event and the measurement of the change in drag. The evaluation of drag requires determination of the change in orbital elements, and an accurate determination requires a day's observation. An alternative scheme for determining the time delay is to obtain a statistical estimate of the phase relations by measuring the

cross spectrums between the drag and radio flux (see, for example, Munk and MacDonald (1960) for a discussion of cross-spectral analysis). The phase angle at a given frequency then fixes the average time interval between the arrival of the decimeter radiation and the change in drag.

Measurements of the cross spectrums show that at the 27-day period the drag on Echo 1 lags behind the 20-centimeter flux by 32°; the drag on Explorer 9 lags by 31°. Thus at the 27-day period there is a 56-hour delay between the emission of the solar radio flux and the response of the atmosphere.

Neither set of statistics is particularly good, though those from Explorer 9 are much better than those from Echo 1. Goodman (1957) shows that there is a probability p that the true phase will lie within the limits $\theta - \Delta\theta$ and $\theta + \Delta\theta$ where

$$\sin^2 \Delta\theta = \frac{1 - R^2}{R^2} \left[(1 - p)^{-2/v} - 1\right]$$

and where R is the coherence and v the number of degrees of freedom. For Explorer 9 there are 8 degrees of freedom, so that there is one chance in five that the true phase lies outside the interval of 20° to 40°, and there is less than one chance in a hundred that the true phase is zero, provided that the errors are normally distributed.

If no allowance is made for the response time of the atmosphere, the 56-hour delay corresponds to a straight line velocity from the sun of 750 km sec^{-1}. Observations of the solar wind made on Mariner 2 (Neugebauer and Snyder, 1962) show a quiet-time average velocity during the interval August 29, 1962, to October 31, 1962, of 560 to 690 km sec^{-1}.

The phase lag determinations thus establish that the 27-day fluctuations in atmospheric drag are associated with the plasma issuing from the sun. The long delay time rules out both ultraviolet radiation and high-energy particles traveling directly from the sun as the principal causes of the 27-day variation. The relative weakness of the statistical data is strengthened by the agreement between the two satellites. The identification of the 27-day fluctuation in drag as primarily due to interaction of the upper atmosphere with the solar wind is important in later considerations of heat sources in the upper atmosphere.

A similar analysis comparing drag on Echo and Explorer 9 with the variation of geomagnetic K and A numbers shows that there is a much weaker correlation with geomagnetic activity.

The spectrum of drag variations shows power both at lower and higher periods than the 27-day variation. No single satellite has

been up long enough to determine in detail long-period variations of the solar atmosphere: it is clear that the 11-year solar cycle does show up in the observations of satellites (Harris and Priester, 1962b; King-Hele and Rees, 1962). Again, the density in the upper atmosphere is correlated with the long-period changes in radio flux.

Construction of Numerical Models

Satellite orbits provide abundant information about the density and its variations in the high atmosphere. However, there is as yet very little information about the chemical composition. Unless the chemical composition is assumed, the density values cannot be converted into temperature, and it is temperature that is of interest in problems of atmospheric escape.

The usual procedure is to compute a model of the atmosphere based on diffusive equilibrium above a certain level. The model can then be tested by comparing its predicted density with the densities derived from satellite determinations. There are many difficulties hampering such a procedure. The nature and distribution of the heat sources are not well determined. It is not clear where the region of convective equilibrium ends and diffusive equilibrium begins. As yet, satellite investigations have not probed in detail the region between 100 and 200 km: yet it is in this altitude range that the atmospheric parameters undergo the greatest change. As a result,

values must be extrapolated through this region. Many of the parameters needed to describe the processes are based on incomplete laboratory data or on uncertain theoretical predictions. Despite these many difficulties, model atmospheres have been constructed that agree well with the densities and the variations in densities derived from satellite observations.

Before reviewing possible models for the upper atmosphere, we first consider methods by which models are constructed. Though the physical principles that govern the distribution of the constituents in a planetary atmosphere are well known, it is essential to have in view the numerous parameters that go into establishing the vertical distribution of density, temperature, and composition.

It is usual to assume that the distribution of atmospheric pressure with height (z) is according to the hydrostatic law,

$$dp/dz = - \rho g \qquad (1)$$

where p is the total pressure, g is the acceleration of gravity, and ρ is the density: both pressure and density will be functions of time as well as of altitude; g varies with height. Equation (1) describes a static equilibrium.

The density is a function of the number density or concentration of the constituents n_i and the molecular masses m_i.

$$\rho = \sum_i m_i n_i \tag{2}$$

It is assumed that the atmosphere behaves like a perfect gas with an equation of state of

$$p = nkT \tag{3}$$

where k is Boltzmann's constant (1.38×10^{-16} erg deg^{-1}), and T is the absolute temperature. Total concentration n is

$$n = \sum_i n_i \tag{4}$$

where the summation is over all constituents. From (3) we see that the vertical variation of pressure depends on the local values of pressure, concentration, and temperature and on the gradients of concentration and temperature,

$$\frac{1}{p}\frac{dp}{dz} = \frac{1}{n}\frac{dn}{dz} + \frac{1}{T}\frac{dT}{dz} \tag{5}$$

Combining (5) with (1) we see that the concentration gradient depends on the local values of temperature concentration and the gradient of temperature,

$$\frac{dn}{dz} = - \frac{\sum n_i m_i g}{kT} - \frac{n}{T} \frac{dT}{dz} \tag{6}$$

Combining (6) and (4), we obtain the general law governing the distribution of a given constituent with height in a static atmosphere:

$$\frac{1}{n_i} \frac{dn_i}{dz} = - \frac{m_i g}{kT} - \frac{1}{T} \frac{dT}{dz} \tag{7}$$

Integration of (7) yields

$$n_i(z,t) = \frac{n_i(z_o,t)T(z_o,t)}{T(z,t)} \exp \left[- \int_{z_o}^{z} \frac{m_i g(z)}{kT(z,t)} \, dz \right] \tag{8}$$

where z_o is some reference height usually taken as the height at which diffusive equilibrium begins.

The variation of the gravitational acceleration with height is given by

$$g(z) = g_o(z_o)[R_o/(R_o + z)]^2 \tag{9}$$

where g_o is the value of the gravitational acceleration at a reference height z_o which is at a distance R_o from the center of the earth.

The concentration of a given constituent depends on the local temperature $T(z,t)$ and on an integral involving the temperature

distribution between z_0 and z. The distribution of temperature thus enters critically into the variation in concentration of a given constituent. The concentration n_i (z,t) also depends on the concentration of the constituent at z_0 as well as the temperature at this reference height. The reference height is usually taken to be equal to the height at which diffusive separation commences and below which convective mixing prevails. A problem of importance is the determination of this level and a description of the gradation between convection and diffusion. In general, the concentration and temperature at the reference level will vary with time.

In a time-varying atmosphere, heat can be transported both by conduction and convection. The vertical transport of heat due to the mass motion resulting from changes of density adds a complication, and we will wish to evaluate the contribution of convection to the total thermal transport. In considering the thermal balance, we will neglect the heat flow caused by the concentration gradients (Dufour effect) as we also neglected the flow of matter caused by a temperature gradient in deriving (8). We will return to the effects of thermal diffusion when considering the motion of a minor constituent such as helium or hydrogen through the atmosphere.

The equation for the concentration of entropy s is

$$ds/dt = q/T \qquad (10)$$

where q is the net accession of heat (erg g^{-1} sec^{-1}) (Eckart, 1960). We wish to transform (10) into an equation for temperature. From thermodynamics the vertical gradient of temperature is related to the equilibrium density and entropy gradients by

$$\frac{1}{T}\frac{dT}{dz}\bigg|_{eq} = (\gamma - 1)\left\{\frac{1}{\rho}\frac{d\rho}{dz}\bigg|_{eq} + \frac{1}{nk}\frac{ds}{dz}\bigg|_{eq}\right\} \quad , \tag{11}$$

where γ is the ratio of the specific heats. The ideal gas law has been used in obtaining (11). The equilibrium gradients appearing in (11) are independent of time.

The temperature, density, and entropy can be written as the sum of the equilibrium values plus a time-varying quantity that is assumed to be small. The validity of such perturbation expansion must, of course, be examined critically. In general, the time-varying quantities, T_1, ρ_1, s_1, will be functions of the horizontal coordinates as well as the vertical coordinates,

$$\begin{aligned} T(z,t) &= T_{eq}(z) + T_1(z,t) + \cdots \\ \rho(z,t) &= \rho_{eq}(z) + \rho_1(z,t) + \cdots \\ s(z,t) &= s_{eq}(z) + s_1(z,t) + \cdots \quad . \end{aligned} \tag{12}$$

Again, we will wish to examine the effects of the horizontal gradients on the thermal balance, and it will be shown that these are relatively small. Equations (10), (11), and (12), together with the expression for the conservation of mass,

$$(\partial\rho/\partial t) + (\partial/\partial z)(\rho w) = 0 \tag{13}$$

yield

$$\frac{\partial T_1}{\partial t} + \frac{\partial T_{eq}}{\partial z} w + (\gamma - 1) T_{eq} \frac{\partial w}{\partial z} = \frac{q}{c_v} \tag{14}$$

w is the velocity in the vertical direction, and c_v is the specific heat at constant volume. Equations (13) and (14) are valid only if the motion is strictly one-dimensional along the vertical coordinate.

Accession of Heat

The net accession of heat q is made up of a number of terms. Heat can enter a volume by ordinary thermal conduction. Viscous dissipation associated with the motion can contribute to the net accession. In the upper atmosphere the principal source of heat is that due to radiation from the extreme ultraviolet, and there are radiative cooling effects.

We first consider the heat due to conduction. Heat transported into a unit volume by conduction is given by

$$\rho q_c = (\partial/\partial z)[K(T)(\partial T/\partial z)] \tag{15}$$

where K is the thermal conductivity. The thermal conductivity of a dilute gas is related to the viscosity μ of the fluid by

$$K = f\mu c_v \tag{16}$$

(Chapman and Cowling, 1939), where f takes on a numerical value of about 2.5 for a monatomic gas and about 1.9 for diatomic molecules. The kinetic theory of gases shows that the coefficient of viscosity can be written as

$$\mu = \frac{5\pi}{16}\left(\frac{mkT}{\pi}\right)^{\frac{1}{2}} \frac{1}{\pi d^2} = BT^{\frac{1}{2}} \tag{17}$$

where m is the atomic mass, and d is the atomic radius. Table I lists the values for A where the thermal conductivity is written in the form

$$K = AT^{\frac{1}{2}} \tag{18}$$

In a multicomponent system the total conductivity can be written as

$$K(T) = \left(\sum_i A_i n_i / \sum_i n_i\right) T^{\frac{1}{2}} \tag{19}$$

and a similar expression holds for the viscosity of a mixture.

TABLE I

	A ($erg\ cm^{-1}\ sec^{-1}\ deg^{-3/2}$)	B ($erg\ sec^{-1}\ deg^{-1/2}$)
H	2.1×10^{3}	6.8×10^{-6}
He	9.0×10^{2}	1.1×10^{-5}
O	3.6×10^{2}	1.9×10^{-5}
$O_2 \cdot N_2$	1.8×10^{2}	1.3×10^{-5}

The magnitude of the vertical velocity can be estimated from the conservation of entropy equation specialized to one dimension:

$$\frac{\partial s}{\partial t} + w \frac{\partial s}{\partial z} = \frac{q}{T} \tag{20}$$

In the spirit of the perturbation treatment, the vertical distribution of entropy is

$$\left.\frac{ds}{dz}\right|_{eq} = \frac{1}{T_{eq}} \left(c_p \left.\frac{dT}{dz}\right|_{eq} + g \right) \tag{21}$$

The vertical velocity w for steady motion is

$$w = \frac{q}{c_p} \left(\left.\frac{dT}{dz}\right|_{eq} + \frac{g}{c_p} \right)^{-1} \tag{22}$$

where g/c_p is the adiabatic gradient for an ideal gas. c_p is the heat capacity at constant pressure. For an ideal gas, the heat capacity at constant pressure can be written as

$$\rho c_p = \sum_i k n_i C_i \tag{23}$$

where k is Boltzmann's constant, and C_i is a constant depending on the constituent

$$\begin{aligned} C_i &= 2.5 \quad \text{monatomic molecules} \\ &= 3.5 \quad \text{diatomic molecules} \end{aligned} \tag{24}$$

In interpreting (22) we note that q/c_p is the rate at which the temperature increases if the heating occurs at constant pressure and without motion. If the fluid is initially in hydrostatic equilibrium in a temperature gradient, $dT/dz\big|_{eq}$, and becomes heated, it rises. If the heating occurs at constant pressure and the fluid always remains in hydrostatic equilibrium, the velocity of rise is

$$(q/c_p)(dT/dz\,\big|_{eq})^{-1} \qquad (25)$$

The fluid as it rises encounters lower pressure and expands. The expansion causes a cooling because of the work done on the surroundings by the expansion. In order to remain in hydrostatic equilibrium, the fluid rises with a smaller velocity, and the adiabatic gradient appearing in (22) takes account of the adiabatic cooling.

To obtain a numerical estimate, we consider conditions at 200 km where the density is of the order of 4.0×10^{-13} g cm^{-3} and the density scale height is about 50 km. If 1 erg cm^{-2} is deposited over a scale height every second, then q is $\approx 3 \times 10^5$ erg g^{-1} sec^{-1}. The vertical velocity is then of the order of 3×10^2 cm sec^{-1}. The presence of a temperature gradient will further reduce the maximum vertical velocity. In addition, we will see that the estimate of 1 erg being deposited over the 50-km height range is on the high side and, therefore, the vertical velocity of 6 meters sec^{-1}

provides a maximum limit for this altitude. The satellite observations indicate at most a 2 percent diurnal change in density at this altitude. This corresponds to a vertical displacement of 1 km of the surfaces of equal density. The corresponding average value of the velocity of surfaces of equal density would be of the order of 2 cm sec^{-1}.

Estimates of the vertical velocity permit an estimate of the rate at which heat is deposited by viscous degradation of mechanical energy; the viscous contribution to the accession can then be compared with the heat accession due to conduction. In addition, we will want to compare heat conducted and heat transported by convection.

From dimensional considerations, we see that the rate of accession of heat due to viscous dissipation q is of the order of

$$\omega w^2/\rho L^2 \tag{26}$$

where L is the smallest length scale involved in the motion. Setting L equal to the scale height, we have that

$$q_v \approx \frac{10^{-5}(10^2)^2}{4 \times 10^{-13}(5 \times 10^6)^2} \approx 10^{-2} \text{ erg g}^{-1} \text{ sec}^{-1} \tag{27}$$

Since the rate of accession from other sources is of the order of

10^5 erg g^{-1} sec^{-1}, the contribution of viscous dissipation is completely negligible. This would be true even if a much smaller scale of motion were involved.

The rate of change of temperature due to heat transported by ordinary conduction is given by

$$\frac{1}{\rho c_v} \frac{\partial}{\partial z} \left(K \frac{\partial T}{\partial z} \right) \tag{28}$$

The order of magnitude of the ratio of heat carried by conduction to that transported by convection is

$$K/\rho c_v L w \tag{29}$$

where L can again be taken as equal to the scale height. At 200 km, the temperature is of the order of 1000° K and the total number density is of the order of 10^{10}. The ratio of the heat transported by conduction to that convected is then of the order of 30. Conduction thus outweighs convection by an appreciable amount, though the total heat transported by convection may not be negligible.

Influence of Horizontal Temperature Gradients

In calculating the net accession of heat, account is taken of the vertical flux of heat. The difference in temperature between

day and night hemispheres will lead to a horizontal flow of heat.

The difference in temperature between the dark and sunlit side is of the order of 600°K at 1000 km, corresponding to an average temperature gradient of 3×10^{-7} °K cm^{-1}. The corresponding heat flow is of the order of 1×10^{-2} erg cm^{-2} sec^{-1}. This heat flow is small compared to the vertical heat input, but may measurably affect the details of the temperature distribution, particularly in the upper level where the vertical gradients are only slightly larger than the horizontal gradients. The effect of the horizontal flow of heat will be to reduce the maximum temperatures.

The high thermal conductivity of the upper atmosphere will prevent large permanent temperature gradients that are not maintained by external conditions. The time required for a temperature anomaly having a scale length L to decay is given by

$$t \approx (\rho c_v / K) L^2 \qquad (30)$$

(Carslaw and Jaeger, 1959). At 400 km, a temperature disturbance with a scale length of 1000 km would decay in about 9 hours. At 1000-km height, the same disturbance would decay in a few minutes.

The large horizontal temperature gradients required by the day-night effect will also produce horizontal velocities. For the horizontal mixing to be effective, velocities on the order of 10^4 cm sec^{-1} would be required for the horizontal mass motion to transport appreciable thermal energies. From considerations similar to those that led to an estimation of vertical velocity, we can show that the horizontal velocities are much too small to give rise to any appreciable horizontal thermal transport.

The small time scale for the vertical flow of heat implicit in (30) assures that the atmosphere will respond to changes in thermal input with a time scale of a few hours at most.

Ultraviolet Heat Source

So far we have considered both the magnitudes of conducted and convected heat and the contribution of viscous dissipation to the thermal budget. Convection is minor compared with conduction, and the effects of viscosity can be neglected. The most important heat source is due to the absorption of solar ultraviolet radiation; this heat source can be written as

$$\rho q_{uv} = \sum_{\lambda} \sum_{i} \epsilon_{i\lambda} n_i(z, t) F_\lambda \sigma'_{i\lambda} \exp [- \tau_{i\lambda}(z,t)] \qquad (31)$$

The summations are over all constituents i and over all wavelengths λ. F_λ is the time-dependent incident flux of wavelength λ at the top of the atmosphere. $\sigma'_{i\lambda}$ is the photoionization cross section of the ith constituents at wavelength λ. $\tau_{i\lambda}$ is given by

$$\tau_{i\lambda} = \int_z^\infty \sigma_{i\lambda} \frac{n_i(z,t)}{\cos\,\theta(t)}\, dz \qquad (32)$$

where $\sigma_{i\lambda}$ is the total absorption cross section. θ is the zenith angle of the sun. $\epsilon_{i\lambda}$ is an efficiency factor for the conversion of the ultraviolet absorbed by the ith constituents into heat.

A determination of the ultraviolet heat source requires a measurement of the ultraviolet radiation F_λ outside the earth's atmosphere; the number densities n_i must be calculated for a particular model. Hinteregger and Watanabe (1962) have given the most detailed review of the values of the solar flux outside the atmosphere. Table II is a summary of their results for the energy flux in various wavelength regions.

Molecular oxygen has the lowest ionization potential among all the major constituents of the atmospheric gases. Radiation at wavelengths shorter than 1027 A will ionize O_2. Radiation in the wavelength region 1350-1027 A is important in the formation of the D layer of the ionosphere and through its interaction with the minor

TABLE II

Flux of Solar Radiation at Short Wavelengths

Wavelength Interval in A	Energy Flux in Wavelength Interval (ergs cm^{-2} sec^{-1})
1350 - 1027	6.0
1027 - 911	0.2
911 - 630	0.3
630 - 370	0.3
370 - 280	0.6
280 - 170	0.8
170 - 10	0.3
10 - 1027	2.5

constituent NO. Atomic oxygen is ionized by radiation below 911 A, and molecular nitrogen is ionized below 796 A. The ionization threshold for helium is 504 A.

The other parameters that enter into an expression derived from laboratory observation or from theory are uncertain in various ways. In particular, the efficiency of the conversion of ultraviolet radiation into heat involves a large uncertainty.

Before considering the detailed application of (31), we note that the heating is proportional both to the number density and to the ionizing flux. As the ionizing flux passes through the highest parts of the atmosphere it is lightly attenuated because of the limited number of absorbers. As it enters into somewhat denser atmosphere, the attenuation goes up, and the heating rate will be a maximum for a given wavelength at an intermediate region in the atmosphere. The detailed calculations suggest that most of the heating is between 150 and 300 km. This is also the region where the number density of ionized particles reaches a maximum.

Of interest is the heat source at altitudes well above the region of maximum heat production. Hinteregger and Watanabe (1962) estimate that the heat sources due to the ionization of oxygen, helium, and hydrogen are approximated by

$$\rho q_O \approx 1.4 \times 10^{-17} \epsilon_O n_O \text{ erg cm}^{-3} \text{ sec}^{-1}$$
$$\rho q_{He} \approx 1.4 \times 10^{-18} \epsilon_{He} n_{He} \text{ erg cm}^{-3} \text{ sec}^{-1} \qquad (33)$$
$$\rho q_H \approx 4.0 \times 10^{-17} \epsilon_H n_H \text{ erg cm}^{-3} \text{ sec}^{-1} .$$

In writing (33), we assume that the attenuation by absorption is not sufficient to bring about a first-order change in the incoming radiation. The values listed refer to an overhead sun.

The absorption and ionization coefficients in (31) are still poorly known. Detailed theoretical calculations have been carried out for atomic oxygen by Dalgarno and Parkinson (1960). The cross section is a function of wavelength varying from about 2.5×10^{-18} cm^2 at 900 A to 1.3×10^{-17} cm^2 at 500 A. The ionization cross section for N_2 is most complicated and varies greatly, having an average value of about 1×10^{-17} cm^2. The lack of secure data on the photoionization and absorption cross section is one of the weaknesses in the construction of any model of the upper atmosphere. In addition, the efficiency of transfer of absorbed photons into thermal energy is even more uncertain (Nicolet, 1959, 1960, 1961a). Some of the absorbed solar energy goes into the excitation of atmospheric particles, and that energy may be lost in the form of radiation. The energy that goes into the dissociation of molecules into atoms is lost as far as the upper regions of the ionosphere are concerned since recombination does not occur in these regions,

but rather the atoms must diffuse down into a region where the density is high enough so that three-body collisions take place. The energy that is absorbed in exciting metastable atomic or molecular states is not available if the atom radiates the energy instead of becoming deactivated by collision. The number of possible metastable states is large, and the evaluation of the definite processes is a complicated matter. Hanson and Johnson (1961) estimate that some 15 to 30 percent of the total energy absorbed in the upper regions of the ionosphere goes into heat.

Radiative Loss of Heat

In determining the net accession of heat it is necessary to consider radiation out of the region of interest. The principal infrared radiator in the upper atmosphere is atomic oxygen (Bates, 1951) which radiates at 63 μ according to

$$O(^3P_1) \rightarrow O(^3P_2{}^2) + h\nu(\lambda = 63\mu) \tag{34}$$

The net accession due to radiation in atomic oxygen can be written as

$$\rho q_{Ox} = n_O A_{12} E_1 \left(\frac{W_1 \exp(-E_1/kT)}{W_2 + W_1 \exp(-E_1/kT) + W_O \exp(-E_O/kT)} \right) \tag{35}$$

which is

$$\rho q_{Ox} = \frac{-1.68 \times 10^{-18} e^{-228/T}}{1 + 0.6e^{-228/T} + 0.2e^{-235.3/T}} n_O \tag{36}$$

n_0 is the concentration of atomic oxygen. E_1 is the energy difference between the $^3P_1$ and $^3P_2$ levels of atomic oxygen; E_0 is the energy difference between the $^3P_0$ and $^3P_2$ levels of atomic oxygen; W_0, W_1, W_2 are the statistical weights at the various levels; and A_{12} is the Einstein coefficient for the transition $^3P_1$ to $^3P_2$.

The numerical coefficient multiplying the concentration of oxygen in (36) will be less than 10^{-18} erg cm^{-3} sec^{-1}. For the radiative losses to equal 1 erg sec^{-1} requires about 10^{18} atoms of atomic oxygen. At a concentration of 10^{11} atoms cm^{-3}, a column 100 km thick would be required to radiate 1 erg cm^2 sec^{-1}. As we shall see, the concentration of atomic oxygen in the upper atmosphere is far less than this. However, the infrared radiation from atomic oxygen can contribute substantially to the heat balance in the height interval 120 - 160 km.

Numerical Models of the Upper Atmosphere

Numerous models for the upper atmosphere incorporating the satellite-derived densities and ultraviolet fluxes obtained from rocket measurements have been calculated. The most detailed analysis has been carried out by Harris and Priester (1962a). Combining (8) and (14), Harris and Priester have determined the time variation of a model atmosphere, taking into account the gain of heat through ultraviolet flux, the loss through radiation by atomic oxygen, and approximating the vertical motion within the atmosphere. In this

way they derive the daily and solar-cycle variation (Harris and Priester, 1962b) of the density concentration and temperature of the upper atmosphere. We present a review of their work in order to illustrate certain of the problems inherent in such calculations and further to indicate the current estimates of the constitution of the upper atmosphere.

Boundary conditions at 120 km. The simultaneous integration of (8) and (14) requires boundary values of the number density of the relevant constituents and the temperature at the level at which diffusive equilibrium replaces convective mixing. The two problems are the determination of the height of this level and the estimation of the constitution of the atmosphere at this level.

Early attempts to determine the composition and height of the diffusion level through the use of mass spectrometers were made by Townsend et al. (1954). During the IGY, successful flights through the critical regions were made (Townsend and Meadows, 1958). The method of determining the region of diffusive separation involved the determination of the argon to N_2 ratio. These flights and Russian flights (Pokhunkov, 1962) showed a decrease of A/N_2 ratio with altitude compatible with the onset of diffusive separation in the 100 to 120-km region.

An alternative method of determining the level at which diffusive separation sets in is through the use of artificial injections of alkali metals by rockets. Blamont and deJager (1961), through the study of alkali vapor trails, show that below about 100 km motion is intense with a strong shearing. Above 100-105 km the alkali vapors diffuse into the atmosphere. A remarkable feature of the observations is that the transitional region is very sharp. There are seasonal variations in the intensity of the motion and also in the height of the transition layer. Since the alkali-vapor technique is limited to the twilight and dawn hours, it is impossible to investigate the daily variations in the height of the transitional layer. Observations on meteor trails (Greenhow, 1959) show a similar transition though the upper height of the turbulent motion has been found at altitudes greater than 100 km.

A major advance has been the determination of the O/O_2 ratio in the region 100 - 135 km by Schaeffer (1963). Schaeffer ejected the spectrometer from the rocket to minimize the effects of rocket exhaust. The instrument was designed to minimize the probability of surface recombination of atomic oxygen. Schaeffer finds at 127 km a ratio of O/O_2 equal to two. The values are in good agreement with the ultraviolet determination of the O/O_2 ratio (Byram et al., 1955; Kupperian et al., 1960) but disagree markedly with earlier mass spectrometer results.

Table III lists the values of the constituents at 120-km altitude used by Harris and Priester and derived from the study of Nicolet (1961a). The structure of the atmosphere depends particularly on the ratio adopted for atomic to molecular oxygen. As has been noted above, atomic oxygen is a radiative cooling agent, and if a high abundance of atomic oxygen is used the maximum temperatures obtained are less. Harris and Priester estimate that a change of a factor of 10 in the ratio O/O_2 decreases the maximum temperature reached at high altitudes by about 250°. The ratio used is consistent with Schaeffer's (1963) measurements.

Required heat sources. Using the above boundary conditions, Harris and Priester (1962a) carried out detailed calculations of the temperature and composition for various heat sources. A remarkable feature of their study was that an ultraviolet heat source symmetrical in time about local noon, vanishing before 6 o'clock in the morning and 6 in the evening, does not yield density variations compatible with satellite observations. The peak value of the heat source depends on the thermal efficiency ϵ. Harris and Priester assigned a thermal efficiency of 0.37 and used Hinteregger's estimate of 2.5 erg cm^{-2} sec^{-1} for wavelengths less than 911 A. The observed response of the atmosphere, however, demands an additional heat source having a maximum in the midmorning, a minimum in the early afternoon, a small amount of heating during the night, and an

TABLE III

Concentration of Constituents at 120 Kilometers

Constituent	Number Density (cm^{-3})
N_2	5.80×10^{11}
O_2	1.20×10^{11}
O	7.60×10^{10}
He	2.50×10^{7}
H	4.36×10^{4}

average magnitude comparable to that provided by the ultraviolet flux. Only if such a heat source is assumed can the calculated densities, assuming the time-independent boundary values given in Table III, be brought into agreement with the observed variations in density derived from satellite observations. The critical point is that the discrepancy is in the phase of the heat source rather than its absolute magnitude. If the magnitude were off, the discrepancy could be attributed to incorrect values for the thermal efficiency or for the ionization cross sections. The phase relations cannot be explained in such a fashion. The model does not take into account the depletion of the neutral population through ionization. The percentage population of the ionized particles is so small that the effects associated with the ionized particles can be neglected. However, there is a possibility of an additional heat source lagging behind the maximum in the solar flux due to the nighttime recombination.

Numerical models. The basic parameter produced by the models that can be compared directly with satellite observations is the density and the variation of density with height. We recall that the absolute values of the density derived from satellites are uncertain, but the variation with time is much clearer. Figures 5 and 6 show the variation of the number density of the neutrals between 0 and 12 hours and 6 and 18 hours. The total number densities are useful in association with Figures 7 through 10, which give the

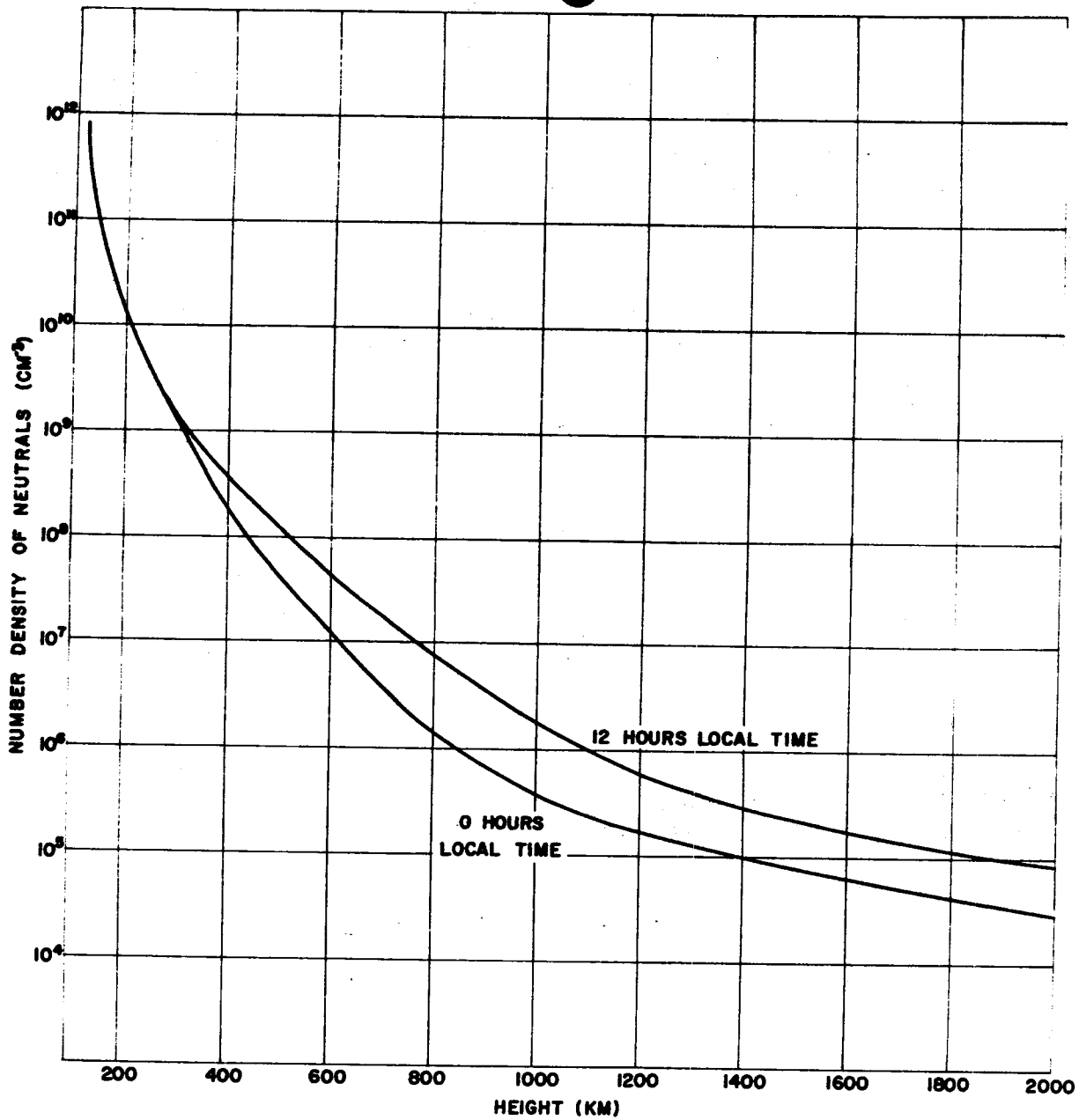


Figure 5: The number density of neutral molecules at 0 and 12 hours local time as a function of altitude. The number densities, derived from the models of Harris and Priester (1962a), correspond to solar conditions during the years 1956 and 1959 during the last solar cycle.

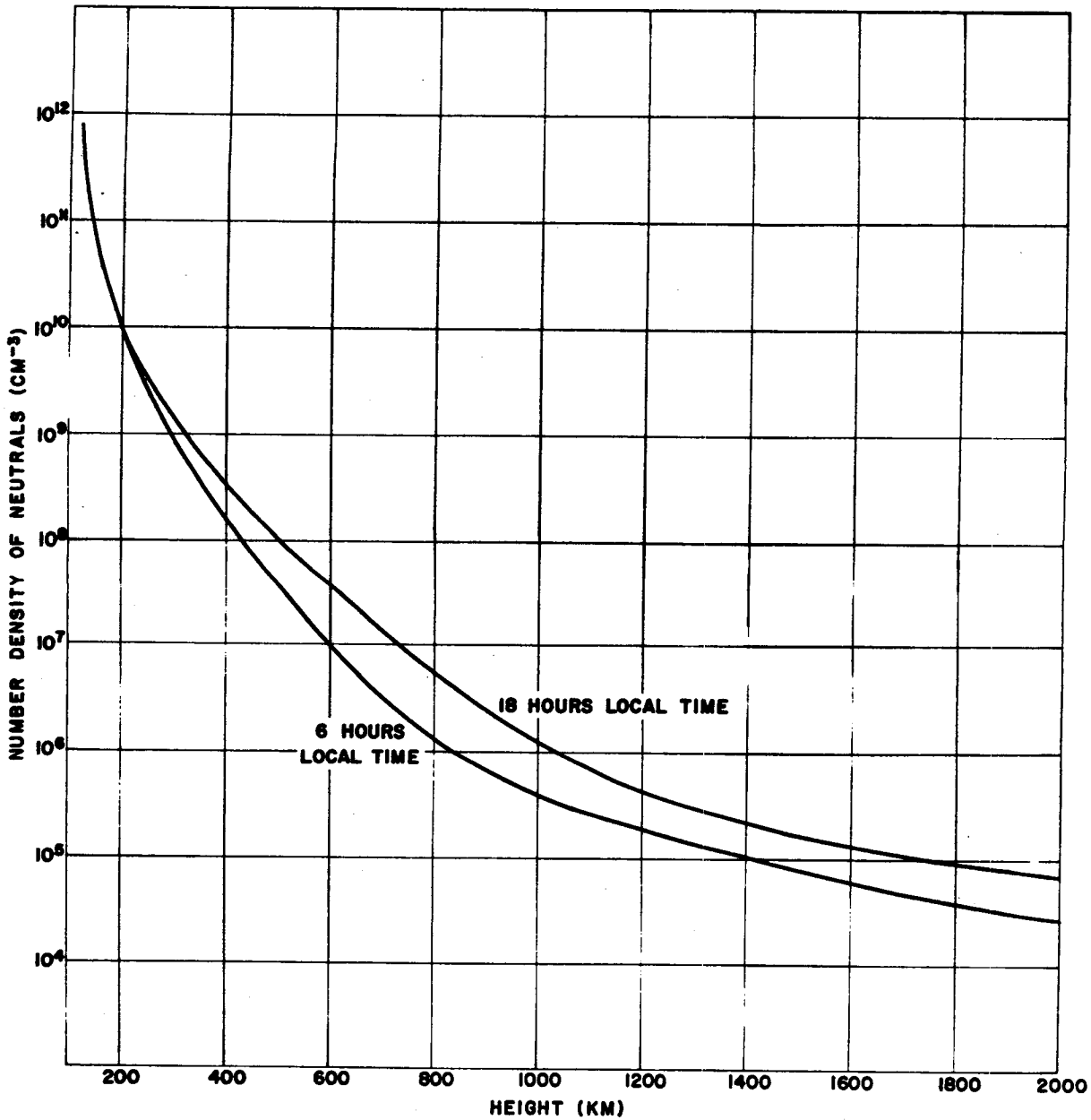


Figure 6: The number density of neutral particles in the high atmosphere at 6 and 18 hours local time.

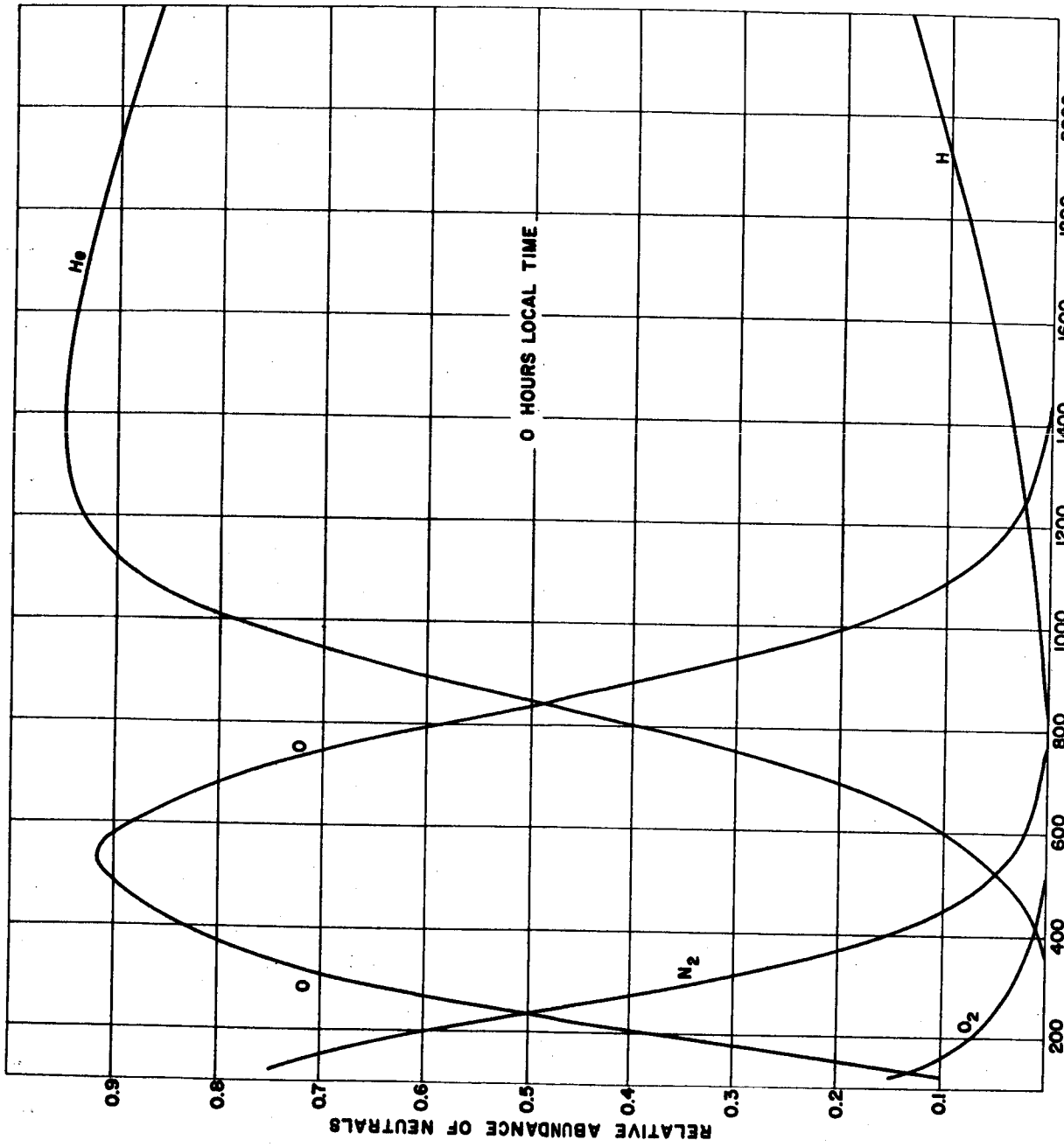


Figure 7: The relative abundance of neutral particles as a function of altitude at 0 hours local time. The absolute abundances can be derived by using Figure 5.

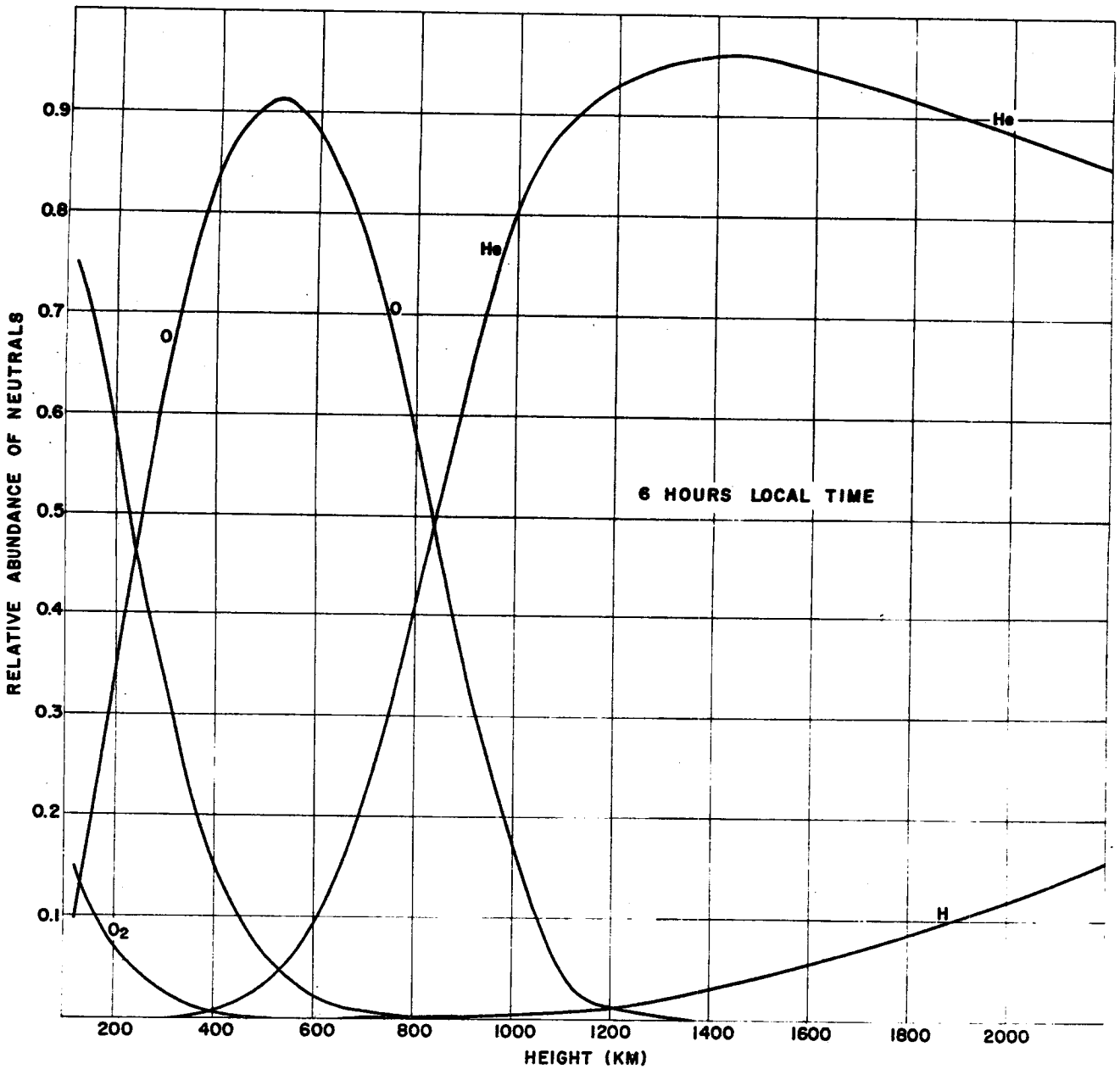


Figure 8: The relative abundance of neutral particles at 6 hours local time. The absolute abundance can be derived by using Figure 6.

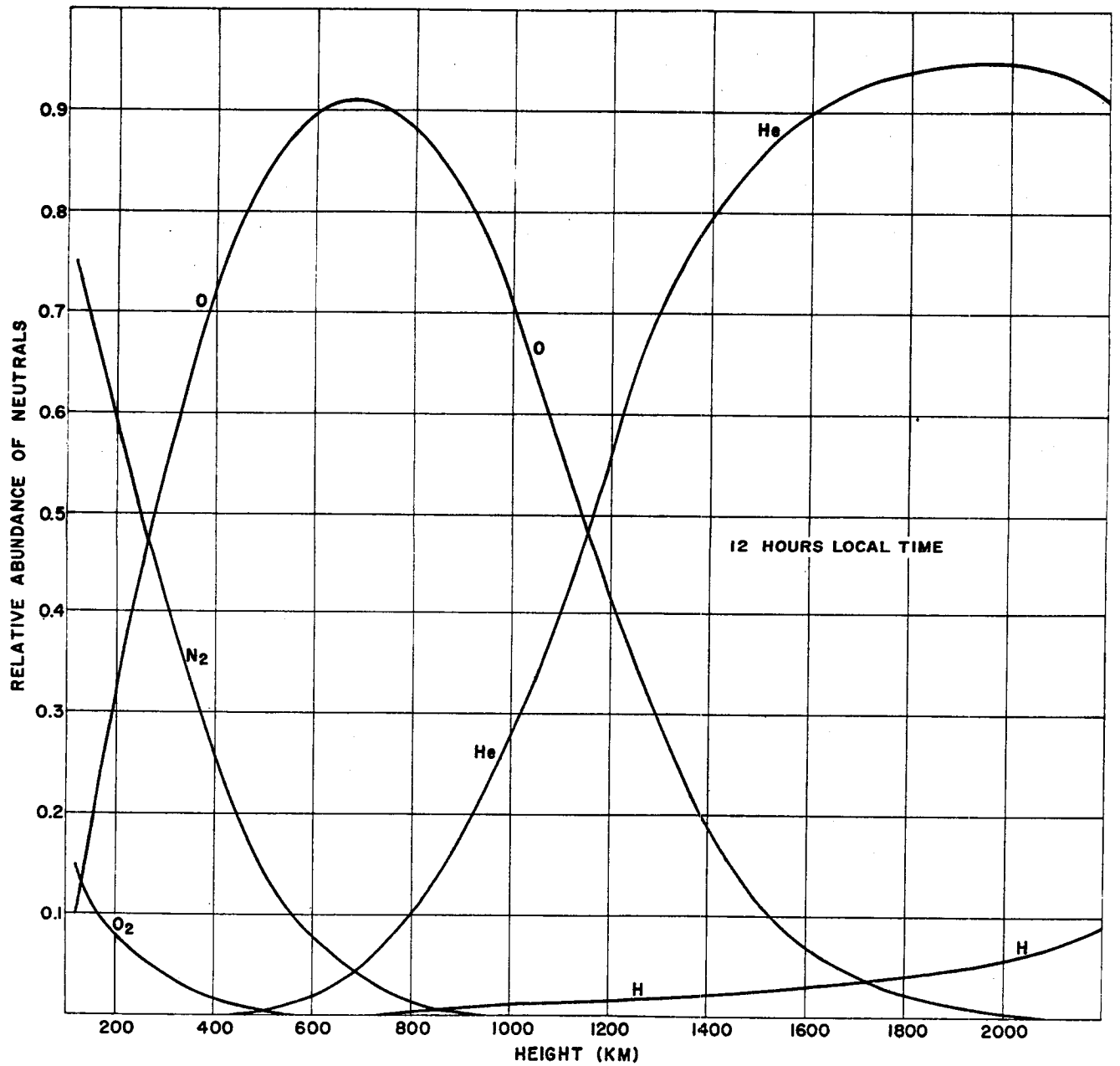


Figure 9: A relative abundance of neutral particles at 12 hours local time. The absolute abundances can be derived by using Figure 5.

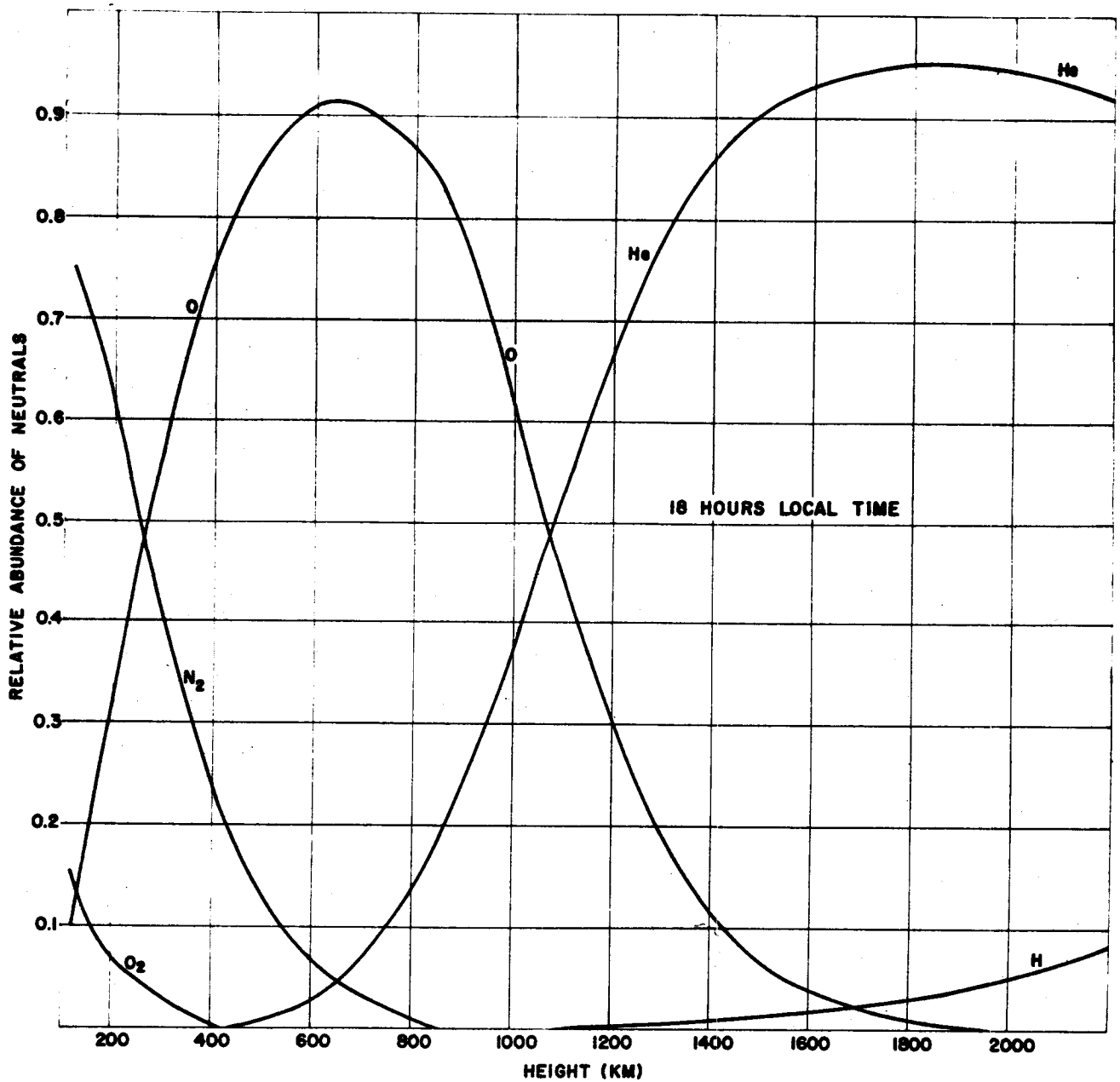


Figure 10: A relative abundance of neutral particles at 18 hours local time. The absolute abundances can be derived by using the number densities given in Figure 6.

relative proportion of the constituent atoms within the upper atmosphere.

Figures 5 and 6 show that the major fluctuations in density take place at altitudes above 300 km. At 1000 km, the maximum change in density is about a factor of 8.

Figures 7 through 10 show the variation of the relative proportions of the major constituents. Three distinct regions in the atmosphere between 120 and 2000 km can be outlined. At noon, between 120 and 250 km, molecular nitrogen dominates. Between 250 and 1150 km, atomic oxygen is the principal constituent, and above 1150 km, helium dominates. Nicolet (1961b,c) first showed that the slow decrease in atmospheric density with altitude between 750 and 1500 km as derived from satellite observations indicated that helium was the major constituent in the region concerned.

Nicolet's prediction is strikingly confirmed by observations of Bourdeau et al. (1962) who found that at an altitude of 1630 km the He^+ to O^+ ratio was of the order of unity. The ratio of the ionized particles cannot be directly converted into ratios of the neutral particles, but the observations do indicate the importance of helium in the upper atmosphere. Hanson (1962) also showed that an ion probe experiment by Hale (1961) can be interpreted as

indicating that the helium atoms are indeed present as suggested by Nicolet. The composition of the atmosphere derived by Hanson is in general agreement with those obtained by Harris and Priester.

The variations in composition during the day-night cycle are illustrated in Figures 7 through 10. During the nighttime, the peak concentration of oxygen is at a lower level than during the day. As the atmosphere expands because of the heating, the maximum in the concentration of both oxygen and helium moves outward.

The variation with temperature, with time, and with altitude in the model atmosphere is shown in Figures 11 and 12. The temperature at midnight is about 1200° at heights above 400 km; the temperature at noon is on the order of 1715° during 1956 and 1959 of the last solar cycle.

Solar-cycle variations. Harris and Priester (1962b) have attempted to estimate the variation in the parameters of the atmosphere during the solar cycle. The basis of their estimates is in the correlation of the postulated heat source with the long-term averages of the 10.7-cm solar flux. The derived densities are then compared with the densities obtained from satellites for the period 1958 - 1961. The models discussed above and represented in Figures 5 through 12 correspond to the years 1956 and 1959 during the last solar cycle. The peak in the solar cycle would be in the period

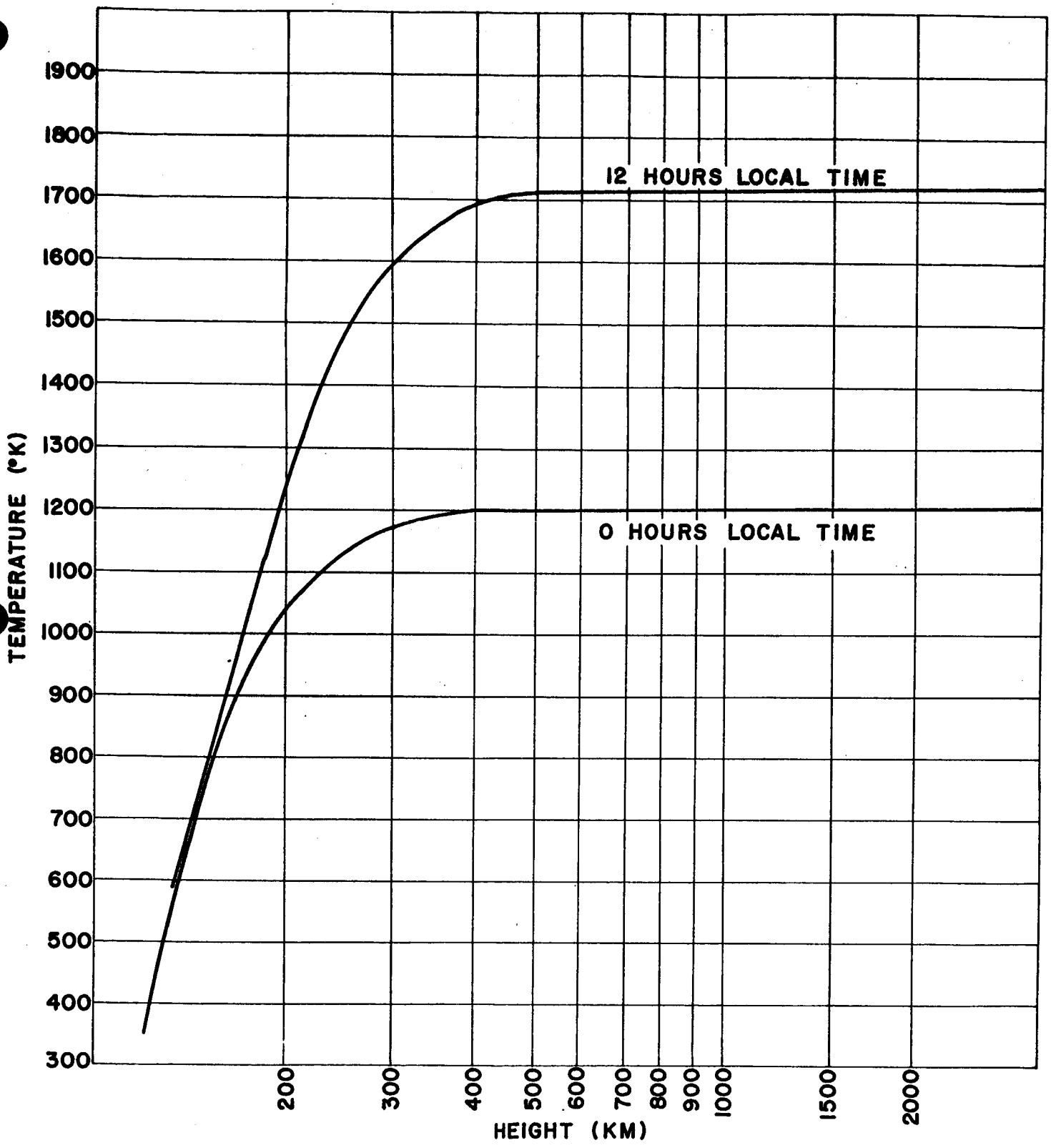


Figure 11: The temperature at 0 and 12 hours local time. The temperature corresponds to conditions during 1956 and 1959 of the last solar cycle (Harris and Priester, 1962a).

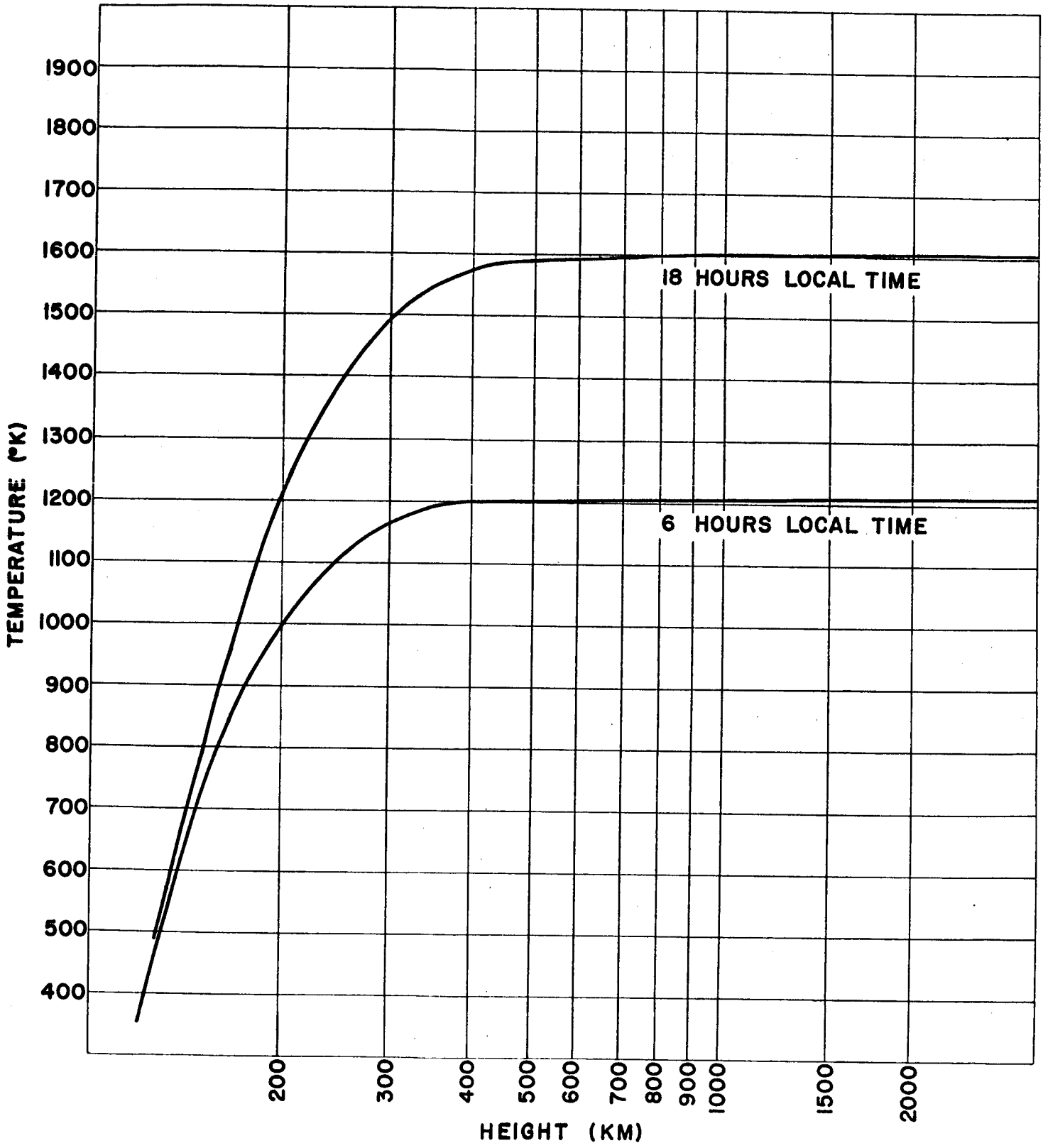


Figure 12: Temperatures at 6 and 18 hours local time. Temperatures correspond to solar conditions prevailing during 1956 and 1959 of the last solar cycle.

1957 - 1958. The maximum, mean, and minimum daily temperatures obtained by Harris and Priester for solar maximum and solar minimum are shown in Table IV. For comparison, King-Hele and Rees (1962) obtain, by using a somewhat simplified model of the atmosphere, an average maximum temperature in 1958 of 1600°; in 1959, 1350°; in 1960 - 1961, 1150°. Again, the agreement with the estimates of Harris and Priester is quite good.

Harris and Priester have assumed that the conditions at 120 km do not vary either daily or over the solar cycle. The time-dependence of conditions at 100 - 120 km is an important problem on which there is yet little evidence. There is no doubt that the O/O_2 ratio varies, probably by a factor of 3 to 4, over the day-night cycle. Such a variation complicates the heat source distribution, since the abundance of oxygen determines the radiative cooling.

Summary of Atmospheric Constitution

The principal results of the satellite studies, when combined with theoretical investigations, show that the temperature of the upper atmosphere varies on a daily basis and with the solar cycle. The details of this variation are still somewhat uncertain, and the absolute temperature assigned at any given time may be in error by as much as 100°. During the maximum of the solar cycle, the daily maximum temperature is of the order of 2150°. During the

TABLE IV

Maximum, Mean, and Minimum Daily Temperatures

Period	(^{o}K) Minimum	(^{o}K) Mean	(^{o}K) Maximum
Solar Maximum	1400	1775	2150
Solar Minimum	590	730	870

minimum of the solar cycle, the maximum temperature may be only 900°. During most of the solar cycle, the maximum daily temperature will be no more than some 1700°. These conclusions have a profound effect as far as any consideration of escape of helium from the upper atmosphere is concerned.

There is an uncertainty about the detailed nature of the composition of the upper atmosphere and the variation of composition with the solar cycle and diurnally. However, it appears that the atmosphere between 200 and 1000 km is, on the average, dominated by atomic oxygen. At higher altitudes, helium is the principal constituent, and at still greater altitudes, perhaps of the order of 4000 km, hydrogen becomes the principal constituent.

The major uncertainties in the construction of a model for the upper atmosphere are concerned both with the nature of the heat source and conditions prevailing at 100 - 120 km. Detailed investigations such as those carried out by Harris and Priester indicate that the ultraviolet heat source is inadequate to account for the phase of the observed diurnal fluctuations. An additional heat source, peaking in the midmorning and at a minimum in the midafternoon, is required to supplement the ultraviolet heat source, which is a maximum at local noon. Since it is the phase of the heat source that is involved, possible underestimates of the thermal efficiency or

errors in the photoionization cross section cannot account for the observed density fluctuations in the atmosphere. However, possible variations in composition at 120 km may account for the phase relations; these effects have not been investigated largely because of the lack of data on the atmosphere in the 100- to 200- km interval. Harris and Priester suggest that high-energy particles may provide this additional energy source. We will later return to the detailed consideration of possible heat sources and their long-term variations.

III. FLUX OF CONSTITUENTS INTO THE ATMOSPHERE

Introduction

In considering the escape of gases from the earth's atmosphere, it is necessary to estimate the rate of influx of these gases into the atmosphere. We will be principally concerned with the two isotopes of helium, He^4 and He^3. However, it is necessary to consider the influx of argon, since knowledge of the argon flux permits an estimate of the flux of helium from the crust into the atmosphere (see Birch, 1951; Damon and Kulp, 1958; Turekian, 1959). The atmospheric content of neon is also of interest, since it sets a limit to the amount of gas that has been accreted from outer space.

Estimation of the rate of release of radiogenic helium and argon from the crust is very complex. It may be that not all the relevant factors have been taken into account. It is important that this point be appreciated, since the conclusion that is to be reached is that the rate of release of helium into the atmosphere at present greatly exceeds the rate at which helium is presently escaping.

Advances in knowledge of the influx of important constituents into the atmosphere have come from two directions. The studies on the abundances of gases in natural waters permit a new method of estimating the influx. Until recently, the influx has followed from estimates based only on the measured heat flows and ages of rocks. In addition, advances in the understanding of the interaction of the

magnetosphere with the interplanetary medium and measurements of the particle fluxes in solar flares allow an estimate of the rate at which certain elements are being accreted.

The content of the various constituents in the atmosphere in particles cm^{-2} is given in Table V. These values are derived primarily from the discussion of Glueckauf (1951) (see also Nicolet, 1957; Urey, 1959).

The Flux of Helium 4

Argon 40 is produced by the radioactive decay of potassium 40. Helium 4 results from the radioactive decay of both uranium and thorium. A rough estimate of the maximum amount of helium and argon that can have entered the atmosphere follows from general considerations of the heat balance of the earth (Nicolet, 1957; Birch, 1951). Potassium, uranium, and thorium are the principal radioactive elements and certainly produce most of the heat that is now reaching the surface of the earth. The present heat flow is 62.8 ± 4.3 ergs cm^{-2} sec^{-1} (MacDonald, unpublished). If the earth had the composition of chondritic meteorites (1.1×10^{-8} g g^{-1} of uranium, 4.4×10^{-8} g g^{-1} of thorium, and 8.0×10^{-4} g g^{-1} of potassium), the surface heat flow would be 59.4 ergs cm^{-2} sec^{-1} if all heat being produced reached the surface. If only the mantle had chondritic composition, then the equilibrium surface heat flow would be 40.1 ergs cm^{-2} sec^{-1}. Several factors contribute to a present heat flow greater than the

TABLE V

Concentration of Rare Gases in the Earth's Atmosphere

Gas	Concentration at sea level, particles cm^{-2}
He^3	1.41×10^{14}
He^4	1.13×10^{20}
Ne	3.84×10^{20}
A	1.93×10^{23}
Kr	2.0×10^{19}
Xe	1.6×10^{18}

equilibrium heat flow. These include the higher rate of heat production in the past, the long thermal time constant of the earth, and the contributions from initial heat (MacDonald, 1959). The close agreement between the equilibrium heat flow from a chondritic earth and the observed heat flow is usually taken as evidence for the chondritic composition of the earth as far as the radioactive elements are concerned.

If, indeed, the earth has a chondritic composition, and if the gases produced in the radioactive decay of potassium and argon are reaching the surface, then the present flux into the atmosphere should be 2.6×10^{6} He^{4} cm^{-2} sec^{-1} and 3.24×10^{6} A cm^{-2} sec^{-1}. The relatively short half-life of potassium 40 means that 4.5×10^{9} years ago argon was produced at approximately 10 times the present rate, so that the average flux of argon over geologic time would be 1.3×10^{7} A cm^{-2} sec^{-1}. If all the argon produced in a chondritic earth had reached the atmosphere, the total content of argon would be 1.8×10^{24} A cm^{-2}. This is a factor of 10 greater than the observed abundance. It is, therefore, clear that not all the argon that has been produced has reached the surface. The validity of the last assertion depends on the potassium abundance of the earth. However, it is extremely unlikely that a consistent thermal model of the earth can be constructed with a potassium concentration one-tenth of that of chondritic meteorites.

The ratio of argon produced to argon in the atmosphere, 0.9, permits an estimate of the earth's present retentivity of argon, provided that the outgassing process has been a continuous one. If 90 percent of the helium is also retained within the earth, then the current flux of helium to the atmosphere would be 2.6×10^5 cm^{-2} sec^{-1}. In a sense, this is a lower limit to the flux of helium for an earth having a chondritic composition, since helium is retained within rocks to a far lesser extent than is argon.

The estimates of the flux of argon and helium into the atmosphere suffer from the fact that they are based on compositional models for the earth which, though reasonable, are still uncertain. Wasserburg et al. (1963) have attempted to obtain an estimate based on the observed content of argon in the atmosphere and on the ratios of helium to argon in natural gases. Measurement of natural gases (Zartman et al., 1961) show that the helium to argon ratio varies from 2 to 200, a value of about 10 being the most frequent. On this basis, Wasserburg et al. conclude that the ratio of the present flux of helium to argon lies between 2 and 20. The present flux of argon is estimated to lie between 7×10^4 and 3.5×10^5 atoms cm^{-2} sec^{-1}, where the factor of 2 uncertainty arises from the possibility that half of the argon present in the atmosphere is primeval. The estimated flux of helium obtained by Wasserburg et al. is then between 1.4×10^5 and 7×10^6 atoms cm^{-2} sec^{-1}. Assuming that all the argon in the atmosphere originated from outgassing, the most frequent ratio

of helium to argon in natural gases would yield a present flux of 3.5×10^6 atoms cm^{-2} sec^{-1}. This estimate lies close to the estimate derived from the arguments based on the present thermal conditions.

The Flux of Helium 3

The two possible sources of helium 3 in the atmosphere are direct production resulting from the interaction with cosmic rays and the injection of helium 3 into the atmosphere from the sun.

A detailed analysis by Fireman (1953) shows that helium 3 may be produced by collision of cosmic-ray neutrons with nitrogen leading to the formation of tritium which then decays to helium 3. Tritium can also be produced by direct ejection in cosmic-ray stars, as can helium 3. Numerous other studies have been carried out evaluating the production of tritium by galactic cosmic rays and the spalation production of helium 3 (Craig and Lal, 1961; Wilson and Fergusson, 1960; Begemann and Libby, 1957; Dostrovsky et al., 1960). These estimates indicate that the galactic cosmic rays produce on the order of 2 He^3 atoms cm^{-2} sec^{-1}.

Direct evidence of the accretion of tritium and helium 3 from the sun was obtained by Fireman et al. (1961) and Schaeffer and Zahringer (1962) from a study of the tritium and helium 3 in the casing of the earth satellite, Discoverer 17, which was flown after the solar flare of November 12, 1960. The amounts of tritium and

helium 3 recovered from the satellite are 2 orders of magnitude too large to be accounted for by spalation of the casing material. The results show that the solar flare radiation intercepted by the satellite consisted of 0.4 percent tritium and 10 percent helium 3. The average yearly integrated intensity of solar cosmic rays with energies above 30 Mev for the years 1956 - 1961 is about 1.1×10^2 particles cm^{-2} sec^{-1} (McDonald, 1963). Thus, the sun may inject on the average some 10 He^3 cm^{-2} sec^{-1} (Flamm et al., 1962).

The accretion rate remains uncertain, since Biswas et al. (1962) did not report helium 3 in emulsions flown shortly after the November 12 flare. Three hundred tracks were examined carefully for the presence of deuterons and tritions. In the energy range accepted by the emulsion, some 3 or 4 particles should have been helium 3 if the ratios obtained from the analysis of the Discoverer satellite are correct.

The Accretion of Neon

The abundance of neon is important, since it sets an upper limit to the gases that may have been accreted from the sun. Neon, like argon, is too massive to have escaped from the earth's atmosphere unless the upper atmosphere temperatures were an order of magnitude greater than they are today. The abundance of neon in the sun's atmosphere is not known directly from spectroscopic observations. Biswas et al. (1962, 1963) obtained estimates of the ratio of neon to

hydrogen and helium in solar cosmic rays. The hydrogen to neon ratio is 1×10^4 and the helium to neon ratio is 6.7×10^2. The numbers are uncertain because of the statistics, but the results are consistent with estimates of cosmic abundances.

The solar cosmic-ray flux of 100 cm^{-2} sec^{-1} is far too low to supply the neon now present in the atmosphere if the rate were maintained over geologic time. A second possible source is the low-energy plasma streaming from the sun. Measurements of the flux (Neugebauer and Snyder, 1962) yield a quiet-time value of 1×10^8 particles cm^{-2} sec^{-1}. If the cosmic-ray abundances are the same for the low-energy flux, then this corresponds to a flux density of 1×10^4 neon atoms cm^{-2} sec^{-1}. The total amount of neon that would be accreted, assuming constant flux over geologic time, would be 1.4×10^{21}, or a factor of 3 larger than the observed abundance. As will be discussed later, the magnetic field effectively shields the atmosphere from the solar wind and, as a result, only when the field is much weaker than it is at present will solar plasma penetrate into the atmosphere. The neon content of the atmosphere could be accreted if the field were much weaker than its present value on the average of about one-third of geologic time.

The maximum amount of accreted hydrogen and helium can be calculated by means of the present abundance of neon and the observed ratios of neon atoms to the lighter isotopes. The resulting maximum fluxes are 2.7×10^7

H cm^{-2} sec^{-1} and 7.2×10^5 He cm^{-2} sec^{-1}. The validity of these maximum figures depends on the assumption that all the neon has been introduced into the atmosphere.

Summary of Flux of Gases into Atmosphere

Table VI summarizes the present information regarding the flux of helium 3, helium 4, and argon into the atmosphere. The range of values permitted for the influx of helium 4 from space results from the uncertainty in the proportion of the solar wind that can penetrate the magnetosphere. If only solar cosmic rays contribute to the upper atmosphere, then the lower values are appropriate. If a substantial fraction of the solar wind does indeed penetrate into the magnetosphere, then a higher value is required.

Using the preferred values given in Table VI, we find that the present abundance of helium 4 would accumulate in 1.3×10^6 years. Only 3.7×10^5 years would be required to supply the atmosphere with its present content of helium 3 if solar cosmic rays make the contribution to the influx indicated in Table VI. If only galactic cosmic rays contribute to the production of helium 3, a time scale of 2.2×10^6 years is required.

TABLE VI

Flux of Gases into the Atmosphere

Constituent	Flux from Earth, cm^{-2} sec^{-1}		Flux from Space, cm^{-2} sec^{-1}	
	Range	Preferred Value	Range	Preferred Value
He^3	...	...	2 to 12	12
He^4	1.4×10^5 to 7.0×10^6	2.0×10^6	6 to 6×10^6	6
A	7.0×10^4 to 3.5×10^5	2×10^5	...	...

IV. THERMAL ESCAPE OF GASES

Introduction

The density of the atmosphere decreases upward. At some level the mean free path of the constituent particles becomes large in comparison with the scale height, and collisions above this level are infrequent. The region in which collisions are negligible is termed the exosphere. The atmosphere grades into the exosphere, though for purposes of the escape theory it is usual to speak of a given level as the base of the exosphere. Below this level collisions are frequent enough so that the particles assume a Maxwellian distribution. From the base of the exosphere, some particles are ejected upward with velocities less than that required for escape. These particles describe elliptic, ballistic orbits and return to the base of the exosphere. Some fraction of the particles describing ballistic orbits have a velocity greater than the escape velocity. These particles will leave the exosphere in hyperbolic orbits. In addition, there is a component in elliptical orbits circling the planet and not passing through regions where the density is high enough for collisions to take place.

A basic problem in the theory of planetary escape is the calculation of what fraction of atoms have a velocity such that their kinetic energy is greater than the gravitational potential energy. The classical theory of escape has been developed by Jeans (1925) and Lennard-Jones (1923). A discussion that has influenced much of

the later work is given by Spitzer (1949). These models assume an isothermal atmosphere with a gas in equilibrium having a Maxwellian velocity distribution up to the base of the exosphere. Above this level collisions are infrequent enough to be unimportant in slowing down outwardly traveling particles.

The problem of escape is complicated by the fact that, for a minor constituent, the base of the exosphere must be supplied through diffusion from below. If the thermal conditions are such that the escape rate of a particular constituent is large compared with the diffusion rate, then diffusion effectively controls the escape of that constituent. Diffusion thus effectively limits escape during short intervals in which locally the temperature may rise much above the average temperature. The region between 100 and 200 km has a very large temperature gradient, and diffusion through this region is complicated by the phenomenon of thermal diffusion.

Classical Theory of Escape

The two necessary conditions for escape to take place are that the velocity of an outgoing particle is greater than a critical velocity v_c and that the atom with this velocity has a negligible chance of undergoing a collision. The critical velocity is defined by the requirements that the particle's kinetic energy is greater than the gravitational potential:

$$\tfrac{1}{2}mv_c^2(z) = mg(R_o)\frac{R_o^2}{(R_o + z)} \qquad (37)$$

m is the mass of the particle, g is the gravitational acceleration at an arbitrary distance R_o from the center of the earth, and z is the height measured from R_o at which the particle has a velocity v. For a gas in thermal equilibrium, the number of particles having a velocity between v and $v + dv$ in an element of solid angle is

$$4\pi r^2 n\left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(-\frac{mv^2}{2kT}\right) v^2 \sin\theta\, d\theta \qquad (38)$$

where $r = R_o + z$; θ is the angle between the velocity vector and the radial coordinate. The flux dF in particles per square centimeter second follows when we multiply (38) by the radial component of the velocity, $v \cos\theta$.

$$dF = 4\pi r^2 n\left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(-\frac{mv^2}{2kT}\right) v^3 \sin\theta \cos\theta\, d\theta \qquad (39)$$

Integration over the angular coordinate yields the number of molecules crossing the unit area of a hemisphere per unit time and integration over velocities from the critical velocity to the infinite velocity yields the flux of particles having velocities that exceed the local velocity of escape.

$$F(r) = n(r)\left(\frac{kT}{2\pi m}\right)^{\frac{1}{2}}\left\{1 + \frac{mg(r)r}{kT(r)}\right\} \exp\left\{-\frac{mg(r)r}{kT(r)}\right\} \qquad (40)$$

Defining the scale height H(r) by

$$H(r) = kT(r)/mg(r) \tag{41}$$

the flux of escaping particles per unit area per second is

$$F(r) = n(r)\left(\frac{g(r)H(r)}{2\pi}\right)^{\frac{1}{2}}\left\{1 + \frac{r}{H(r)}\right\}\exp\left\{-\frac{r}{H(r)}\right\} \tag{42}$$

The expression for F(r) gives the outward radial flux at a level r of the particles whose total velocity exceeds the critical escape velocity defined by (37). If there are no further collisions at distances greater than r, F(r) equals the escape loss; otherwise it is just the outward flux. The physical interpretation of (42) thus depends critically on the relative frequency of collisions beyond a given level.

The conditions for the validity of (42) have been discussed in great detail by Opik and Singer (1959, 1960), Brandt and Chamberlain (1960), Herring and Kyle (1960), and Aamondt and Case (1962). Within the region of the atmosphere where collisions are frequent, the Maxwellian distribution on which (42) is based holds quite accurately, and (42) gives a direct estimate of the outward flux; however, part of this outward flux is balanced by an inward flux because a certain fraction of the particles in their upward paths collide and return. The net outward flux is then less than that given

by (42). As the base of the exosphere is approached, the total net flux is more closely approximated by (42), but because of the absence of collision the Maxwellian distribution holds less exactly.

It should be noted that the scale height refers to the particular escaping component. For the case where the scale height is much less than the distance to the center of the earth, (42) simplifies to

$$F(r) = n(r)\left(\frac{g(r)}{2\pi}\right)^{\frac{1}{2}} H(r)^{-\frac{1}{2}} \exp\left\{-\frac{\pi r}{H(r)}\right\} \tag{43}$$

For a minor constituent, the escape level r_c can be defined as the height at which

$$\sigma \int_{r_c}^{\infty} n(r)\, dr = \tfrac{1}{2} \tag{44}$$

where n is the total number density, and σ is the effective collision cross section. For the escape of a minor constituent, (44) can be approximated by

$$H_c n_c \sigma = \tfrac{1}{2} \tag{45}$$

where H_c is the scale height at the critical level, and n_c is the number density at the escape level. For an atmosphere dominated by oxygen atoms, σ is about 7×10^{-16} cm^2, so that the condition at escape level is

$$H_c n_c \approx 7 \times 10^{14} \quad . \tag{46}$$

For the models discussed in Section II, at a scale height of 100 km, the number density is of the order of 7×10^7. The escape level is thus of the order of 500 to 600 km, depending on the time of day. It is also to be noted that the escape level is in a region in which the atmosphere is nearly isothermal.

Theory of Diffusion of a Minor Component

We next consider the laws governing the diffusion of a constituent through the atmosphere. As has been noted above, the rate of diffusion may control the rate of escape when the rate of escape is high. The diffusion of one gas through another in the presence of a temperature gradient has been treated in detail by Chapman and Cowling (1939). For simplicity, we assume a mixture of two gases, A and B, where the number density of gas A is much less than the number density of gas B;

$$n_A \ll n_B \quad n = n_A + n_B \approx n_B \tag{47}$$

We denote the flow of particles in the positive z direction across unit area per unit time by F_A. The relation between the flow and the gradients of composition and temperature can be described by linear relationships if the gradients are not too high.

$$F_A(z) = - nD \left\{ \frac{d(n_A/n)}{dz} - \frac{n_A n_B}{n^2 T} \frac{dT}{dz} \right\} \qquad (48)$$

D is the ordinary coefficient of diffusion, and α is the thermal diffusion factor. In general, α is positive. This means that in the steady state the lighter gas tends to concentrate toward the higher temperatures.

Since the concentration of the constituent A is small compared with the bulk composition, the diffusion will not upset the hydrostatic equilibrium of the bulk gas. In view of this, (48) can be rewritten in terms of the vertical concentration of the constituent A as

$$\frac{dn_A}{dz} = - n_A(z) \left[\frac{1}{H_A(z)} + \frac{\{ 1 - \alpha \}}{T} \frac{dT}{dz} \right] - \frac{F_A(z)}{D} \qquad (49)$$

H is the scale height for the major constituent. The first term on the right-hand side of the equation gives the equilibrium distribution for a diffusing component where no account is taken of the sink corresponding to the escaping gas.

If the last term in (49) is neglected, the ratio of the minor constituent A at two levels, z_1, z_2, is given by

$$\frac{n_A(r_1)}{n_A(r_2)} = \left(\frac{T(r_2)}{T(r_1)} \right)^{1 + \alpha - \beta} \left(\frac{n_B(r_1)}{n_B(r_2)} \right)^{\beta} \qquad (50)$$

where β is the ratio of the molecular masses of A and B.

$$\beta = m_A/m_B \tag{51}$$

(Bates and McDowell, 1957).

Equations (49) and (43) can be combined to yield the expression for the vertical distribution of a minor escaping component,

$$\frac{dn_A}{dz} = - n_A(z) \left[\frac{1}{H_A(z)} + \frac{(1 - \alpha)}{T} \frac{dT}{dz} \right] - \frac{n(z)}{D} \left[\frac{g(z)}{2\pi} \right]^{\frac{1}{2}} H_A^{-\frac{1}{2}}(z) \exp \left\{ - \frac{r}{H_A(z)} \right\} . \tag{52}$$

For a two-component gas with collision cross section σ, the coefficient of diffusion D is

$$D = \frac{3}{32} \frac{1}{n\pi a^2} \left(1 + \frac{m_A}{m} \right)^{\frac{1}{2}} \left(\frac{8kT}{\pi m_A} \right)^{\frac{1}{2}} \tag{53}$$

where a is the atomic radius (Chapman and Cowling, 1939). The diffusion coefficient thus varies inversely with the total concentration of the gas and directly as the square root of the temperature.

Equations (50) and (43) can be combined to yield the escape flux at the critical level in terms of the concentration of the minor constituent at the level at which diffusive separation begins,

provided that the back effect of the diffusion on the vertical distribution of the minor constituent is neglected. Alternatively, we can calculate the escape flux for a given component using (43) and the concentrations derived for various model atmospheres discussed in Section II. The differences in the results arise from the fact that thermal diffusion is neglected in calculating the model atmospheres. We note from (50) that the effect of thermal diffusion is to increase the concentration of the minor constituent at the escape level, since α is positive. The proportional increase of the number density of the minor constituent at the escape level due to thermal diffusion is equal to

$$[T(z_o)/T(z_c)]^{-\alpha} \tag{54}$$

where z_o is the height at which diffusive separation begins, and z_c is the critical height. The magnitude of α is uncertain, but it is customary to take values of about 0.4 (Grew and Ibbs, 1952). We can thus calculate the escape flux for a model atmosphere using (43), making the correction for thermal diffusion given in (54). The calculated flux will then be a maximum flux, since no account is taken of the effect of escape on the diffusion. Bates and McDowell (1957) show that the back effect of diffusion is negligible except when the escape rates become large.

Escape of Helium from the Earth's Atmosphere

Using the model for the atmosphere described in Section II, we find that, at a temperature of 1700° at the critical level, the escape rate for helium 4 is 7×10^4 atoms cm^{-2} sec^{-1}. At the same temperature, the escape rate for helium 3 is 6 atoms cm^{-2} sec^{-1}. At 2100°K, the escape rate of helium 4 is 1.2×10^6 particles cm^{-2} sec^{-1}, and the corresponding escape rate of helium 3 is 50 particles cm^{-2} sec^{-1}. The numerical values obtained when we use the model atmosphere described in Section II are somewhat greater than those obtained by Bates and McDowell (1959) because the models assume a higher concentration of helium at the beginning of the diffusive layer.

Consideration of the solar cycle and diurnal fluctuations in temperature at the escape level indicates that the temperature at the escape level is 2100° less than 1 percent of the time. The temperature would be at 1700° or above less than 10 percent of the time. An integration in time over a solar cycle yields an average rate of escape of helium 4 of 6×10^4 particles cm^{-2} sec^{-1}. This rate of escape corresponds to the rate of escape of helium 3 of 4 particles cm^{-2} sec^{-1}. Both these values are less than the estimated rates of influx into the atmosphere. The rate of influx of helium 4 is at least a factor of 2 greater than the indicated escape rate and may be an order of magnitude greater. It is thus clear that helium 4 cannot be in a steady state at present if thermal escape is the

mechanism by which helium 4 leaves the atmosphere. This point was first noted by Bates and McDowell (1957), and newer data fully confirm their suggestions.

Since the present content of helium would accumulate at a time of the order of 2×10^6 years, it is clear that either some mechanism by which helium could escape has been overlooked or that conditions in the past differed from those holding at present. In the second suggestion what is needed is an effective heat source that will raise the average temperature of the escape level to some 2200°.

The additional heat flux needed to raise the average temperature to 2200° would require detailed investigation, since the number of competing effects is large. However, a rough estimate follows from using (14) and substituting numerical values appropriate for the models discussed in Section II. An increase in the total heat flux of 1 erg cm^{-2} sec^{-1} increases the average temperature about $900^\circ K$. The average temperature over the last solar cycle has been about $1400^\circ K$. An additional heat flux of about 1 erg cm^{-2} sec^{-1} is required for helium to escape at a rate sufficient to keep up with the present production. An additional heat source of 2 ergs cm^{-2} sec^{-1} would result in an escape flux of helium 4 greater than 10^7 atoms cm^{-2} sec^{-1}. These estimates must necessarily be crude, since we assume that the conditions at the beginning of the diffusive layer remain constant while the conditions at greater heights change in response to the changing energy input.

At present, helium 3 is also out of equilibrium with a production rate slightly greater than the escape rate if only galactic cosmic rays contribute to the production of helium 3. If, in addition, there is an influx from solar cosmic rays, then the excess of the input over the escape rate is about 10 atoms cm^{-2} sec^{-1}. The content of helium 3 has built up in a time of about 5×10^5 years. Since the content of helium 4 can be built up in a time of about 2×10^6 years, it can be concluded that within the last 10^5 to 10^6 years conditions in the upper atmosphere have been radically different if thermal escape is the loss mechanism.

The comments of the above paragraphs depend on the assumption that thermal processes alone are responsible for the escape of helium. Bates and Patterson (1962) made the important suggestion that the recombination of HeO^+ produces helium atoms with a kinetic energy in excess of the 2.4 electron volts that are required for escape from the exosphere into interplanetary space. A nonthermal process for the escape provides a neat solution to the helium problem, but the relevant rates of reactions and energy relations are most uncertain.

V. EFFECT OF THE MAGNETIC FIELD

Introduction

The direct effects of the magnetic field on escaping particles are well understood. There is no interaction between the magnetic field and the neutral particles. As far as ionized particles are concerned, the magnetic field restricts their motion to a direction tangent to the field, so that ionized particles cannot escape except where the field lines open out into outer space. The indirect effects of the magnetic field are complicated. The magnetic field acts as a shield to the onrushing solar plasma, effectively fending off the low-energy particles and preventing these particles from entering and heating the atmosphere. In addition, the magnetic field also interacts with the plasma; the plasma generates magnetohydrodynamic and low-frequency electromagnetic waves that penetrate into the ionosphere where they are attenuated. Dessler (1959) first pointed out the possible importance of magnetohydrodynamic waves as a heat source in the upper atmosphere. In addition, the interaction of the field with the plasma will result in a disordered electric field. The electric fields may accelerate electrons into the kilovolt range. The energized electrons provide an additional heat source if they can penetrate the magnetic field.

The behavior of ionized particles in the upper atmosphere differs from the behavior of the neutral particles principally because of the long-range Coulomb force interaction between ionized particles and the

interaction of the ionized particles with the magnetic field. As a result of the Coulomb interaction, a charged particle exchanges its momentum not by single collisions, but in small bits through distant interaction. Because of the small but frequent exchanges of momentum, a particle may transfer an appreciable part of its momentum in distances short compared with the distances a particle must travel to exchange all its momementum. The effective collision cross section is thus not easily defined, but for temperatures of the order of 1000^{o} the cross section for the interaction between ions is about 10^{5} as great as the cross section for interaction between neutral particles.

In the exosphere, the gas is partially ionized and there are several mean free paths corresponding to neutral-charged particle collisions, neutral-neutral collisions, and charged particle interactions. In addition, the inelastic processes such as ionization-recombination charge exchange introduce further complications. The difference between the neutral-neutral mean free path and the neutral-charged particle mean free path is small. Over the entire exosphere, the neutral gas mean free path is in excess of 10 km, and is large compared with other lengths except the scale height that characterizes the motion of the charged particles. It is because of this large difference in the representative length scale of neutral and charged particle gas that it is usually assumed that the two gases act independently. However, for large disturbances in the lower exosphere

this no longer holds true, and the neutral gas exerts a heavy influence in damping the organized motion of the ionized gas.

Because of the interaction of the magnetic field with the charged particles, it is possible to treat the ionized gas as a continuum even for processes where the length scales are short compared to the effective mean free path. Because of the magnetic field, the charged particles spiral around the lines of force. The motion of the charged particles in directions at right angles to the line of force is greatly restricted, whereas the motion along the line of force is largely independent of the magnetic field.

Direct Effect of the Field

The magnetic field will exercise a controlling influence on the escape of particles if a particle leaving the exosphere with a velocity greater than the escape velocity becomes ionized before leaving the region dominated by the earth's magnetic field. The escape velocity at 600 km is 10.7 km sec^{-1}. The boundary of the magnetosphere is at about 10 earth radii on the sunlit side (Cahill and Amazeen, 1963). It would thus take an escaping particle less than two hours to pass out of the region where the earth's magnetic field might interact with the particle. The outgoing particle can become ionized either by charge exchange or by photoionization. Assuming an average atmospheric density of 10^3 ionized particles cm^{-3}, and multiplying by the mean velocity and the charge exchange cross

section, we obtain the result that a particle would have to spend about a week in the outer exosphere before becoming ionized. The time scale for ionization due to interaction with extreme ultraviolet is even longer, on the order of a year. It is thus clear that on the average an escaping neutral atom will pass through the magnetosphere without becoming ionized.

Thermal Processes Associated with the Magnetic Field

The indirect effects of the magnetic field on the temperature in the high atmosphere are speculative, and no attempt will be made to present a quantitative theory.

In considering the indirect effect of the magnetic field, we need to review data on the solar wind and the theory of the interaction of the solar wind with the earth's magnetic field. Flights of Explorer 10 and Mariner 2 established certain of the properties of the low-energy particle radiation from the sun. Neugebauer and Snyder (1962), from observations on Mariner 2, found that for quiet, nonstorm conditions the plasma flux was 1.2×10^{8} cm^{-1} sec^{-1}, with an energy density of 4.4×10^{-9} erg cm^{-3} and a magnetic field energy density of 1.0×10^{-10} erg cm^{-3}. For a velocity of 600 km sec^{-1}, the energy flux is then 0.26 erg cm^{-2} sec^{-1}. These values are in general agreement with data obtained by Explorer 10 (Rossi, 1962).

The interaction of the plasma and the earth's magnetic field will lead to a distortion of the field. This problem has been studied in great detail by Beard (1960, 1962), Dungey (1961, 1962), Hurley (1961), and Slutz (1962). In these models, the dynamic pressure of the plasma with n particles of mass m_p and velocity v is $nm_p v^2$, and is balanced by the magnetic pressure $B^2/8\pi$ interior to the bounding surface. The magnetic pressure includes contributions both from the dipole and from the current field associated with the cavity surface.

The calculations of the boundary shape are most appropriate along the sun-earth line, whereas the shape of the magnetic cavity in the downwind direction remains most uncertain. Equating the magnetic and dynamic pressure yields the distance from the center of the earth r_o to the bounding cavity along the sun-earth line:

$$r_o = [\mu_o/32\pi^2]M^{1/3}/p_o^{1/6} \qquad (55)$$

where M is the magnetic moment, and p_o is the dynamic pressure. The linear dimensions of the cavity thus vary as a one-third power of the surface field and inversely as the one-sixth power of the plasma pressure. The interaction cross section of the earth varies as $B^{2/3}$.

Using values for the interplanetary flux and observations of the surface field, we see that the radius of the cavity in the

direction toward the sun is of the order of 9 earth radii (Cahill and Amazeen, 1963). The interaction cross section of the earth with the solar plasma is thus of the order of 300 times the cross-sectional area of the solid earth. If all steady energy of the solar plasma intercepted by the magnetosphere could be funneled into the atmosphere, the energy flux into the atmosphere would be about 100 ergs cm^{-2} sec^{-1}.

Most of the studies of the magnetospheric cavity have been concerned with the steady-state shape. The plasma streaming by undergoes a wide variety of fluctuations (Neugebauer and Snyder, 1962). The observed low-frequency fluctuations in pressure will interact with the bounding surface and give rise to hydromagnetic waves. In addition, higher-frequency fluctuations associated with the disordered electric field will give rise to low-frequency electromagnetic radiation.

The generation and transmission of hydromagnetic waves is a complicated problem (MacDonald, 1961). Under conditions in which the magnetic field pressure is large compared with the thermal pressure, one mode is propagated almost spherically at the Alfvén wave velocity. Another mode propagates the vorticity of the fluid motion one-dimensionally along the magnetic lines of force, again at

Alfvén velocity. The third mode is a perturbed-sound mode and is unimportant for conditions where the mean free path along the line of force is very large.

As the plasma streams past the bounding surface of the magnetosphere at velocities greater than the local Alfvén velocity, waves can be generated in the surface of the magnetosphere either through shear instability or by the fluctuations in the dynamic pressure (Dessler and Fejer, 1963). The action of the fluctuations in dynamic pressure can be visualized in a straightforward fashion. We suppose that there is a pressure fluctuation with a certain length scale L. This fluctuation will tend to produce longitudinal waves of the same scale L traveling at various angles to the solar wind. In general, the wind speed will be different from the phase velocity of the waves. However, if the component of the wind speed, resolved in the direction of the wave, just equals the velocity of the wave, then there is a matching between the fluctuation in pressure and the mode of oscillation. A kind of resonance takes place, and the wave amplitude builds up. The local Alfvén velocity at the bounding surface is of the order of one-tenth the streaming velocity of the plasma. Since the rate of build-up is greatest when the resolved component of the velocity equals the phase velocity, we should expect that the fluctuations in pressure would be most effective on the circular region on the bounding cavity where the circular region subtends an angle of about 10°.

Waves generated along the bounding surface will be transmitted inwardly with a longitudinal mode propagating spherically while the transverse mode is guided along the lines of force. Both modes are trapped within the magnetosphere. These modes then pass into a region where the density of neutral particles becomes so high that the collisions of the ionized particles with the neutral particles remove energy from the wave at a high rate (Dessler, 1959; Francis and Karplus, 1960; Karplus et al., 1962). However, part of the energy can be converted into electromagnetic energy and propagated through the earth's neutral atmosphere as ordinary electromagnetic waves. Observations at ground level can thus provide some estimate of the intensity of hydromagnetic waves in the upper atmosphere. For example, Ness et al. (1962), in a power spectral analysis of geomagnetic fluctuations at Fredericksburg, Virginia, found, in the frequency interval 0 to 0.5 cps, energy of 0.1 erg cm^{-3}. The spectrum peaks at about 0.1 cps with a Q of about 6. Similar values of Q for giant pulsations have been obtained by Sugiura (1961). Since Q is related to the rate of energy dissipation by

$$\overline{dE}/dt = 2E\pi fQ^{-1} \tag{56}$$

where E is the peak energy, and the left-hand side is the energy dissipated averaged over a cycle, the quiet-time rate at which energy is dissipated in the frequency range from 0 to 0.5 cps is then of the order of 1×10^{-14} erg cm^{-3} sec^{-1}. With an Alfvén velocity on

the order of 10^9 cm sec^{-1}, the energy flux into the ionosphere due to hydromagnetic wave dissipation in this frequency range is 10^{-5} erg cm^{-2} sec^{-1}.

The fluctuations in the plasma at the bounding surface contain sufficient energy to account for the extra heating required by the satellite data. It is apparent both from ground observations and from satellite measurements that the energy available in the plasma fluctuations is not radiated into the ionosphere by the hydromagnetic waves, at least in the frequency range of 0 to 1 cps. This raises the intriguing question as to the energy density within the frequency range 1 to 1000 cps. Observations both at ground level and in space are inadequate to give a proper estimate of the energy flux in this range.

VI. SUMMARY AND CONCLUSIONS

The high degree of correlation between the solar activity as measured by the 10.7-cm flux and the change in density of the atmosphere as measured by the satellite drag, coupled with the phase relations, strongly suggests that the solar wind is involved in heating the ionosphere. However, this heating cannot take place at the low-frequency end of the hydromagnetic wave spectrum. The additional energy may be transferred by higher-frequency waves or by high-energy electrons. However, neither mode of energy transfer from the plasma to the ionosphere has definitely been identified.

Until the present additional heat source is clearly identified, considerations of the possible change of the thermal conditions in the upper atmosphere must remain speculative. However, a few comments seem appropriate. If high-frquency hydromagnetic and low-frequency electromagnetic radiation is an appropriate heat source, this heat source would increase with an increase in the cross-sectional area of the magnetosphere and with an increase in the fluctuations in the particles emitted from the sun. The cross-sectional area varies as $B^{2/3}$. An additional heat source in the past averaging 1 erg cm^{-2} sec^{-1} would require that the cross-sectional area should double, and this in turn would require a magnetic field 2.8 times the strength of the present field. This assumes, of course, that the energy in the high-frequency hydromagnetic wave is proportional

to the cross-sectional area. If it is, even the historical fluctuations in the magnetic field must have produced important changes in the temperature of the upper atmosphere.

An alternative suggestion might be that the magnetic field has been much weaker in the past than it is at present. This would then allow the solar flux to enter. The protons would deposit their energy at heights above about 200 km, since the mean energy of the solar wind is of the order of 1 to 10 kev. The energy flux of 0.3 erg cm^{-2} sec^{-1} would be appropriate, since the cross-sectional area for interaction with the plasma would not be enlarged by the magnetic field. In this case, the mean energy of the plasma rather than the fluctuations are used to heat the atmosphere. The present quiet-time flux would not be sufficient to raise the temperature to that required for helium to escape. However, the active-time solar flux may be a factor of 4 to 5 times greater than the quiet-time flux if the solar wind is proportional to the 10.7-cm radial flux. However, this increase in flux would hold only over a short part of the solar cycle, and the temperatures on the average would not be great enough to provide the required escape of helium.

It thus seems clear that the removal of the magnetic field is not sufficient by itself to provide the required escape of helium. An increase of the field by a factor of 2 or 3 would give rise to the required escape flux, provided that the energy in the fluctuating

plasma could be transmitted from the boundary of the magnetosphere and deposited within the ionosphere. Over geologic time, on the average, the field would have to be twice the present field ten percent of the time.

Alternatively, it may be assumed that the ultraviolet part of the solar spectrum undergoes long-period variations that are of the same order as its mean value. This again could account for the loss of helium. One additional source of energy which has not been considered is the radiation of gravity waves from the surface into the atmosphere. Calculations coupled with observation of the ground-level pressure spectrum (MacDonald, unpublished) indicate that this source is not sufficiently strong to produce the required extra heating of the upper atmosphere. Furthermore, it is difficult to imagine, on time scales of the order of a million years, large-scale changes in the source of the gravity waves, since they are generated by motions in the lower atmosphere.

REFERENCES

Aamondt, R. E., and K. M. Case, Density in a simple model of the exosphere, Phys. Fluids, 5, 1019-1021, 1962.

Bates, D. R., The temperature of the upper atmosphere, Proc. Phys. Soc. London, B, 64, 805-821, 1951.

Bates, D. R., and M. R. C. McDowell, Atmospheric helium, J. Atmospheric Terrest. Phys., 11, 200-208, 1957.

Bates, D. R., and M. R. C. McDowell, Escape of helium, J. Atmospheric Terrest. Phys., 16, 393-395, 1959.

Bates, D. R., and T. R. L. Patterson, Helium ions in the upper atmosphere, Planetary Space Sci., 9, 599-605, 1962.

Beard, D. B., The interaction of the terrestrial magnetic field with the solar corpuscular radiation, J. Geophys. Res., 65, 3559-3568, 1960.

Beard, D. B., The interaction of the terrestrial magnetic field with the solar corpuscular radiation, 2, Second-order approximation, J. Geophys. Res., 67, 477-483, 1962.

Begemann, F., and W. F. Libby, Continental water balance, ground water inventory and storage times, surface ocean mixing rates and worldwide water circulation patterns from cosmic-ray and bomb tritium, Geochim. Cosmochim. Acta, 12, 277-296, 1957.

Birch, F., Recent work on the radioactivity of potassium and some related geophysical problems, J. Geophys. Res., 56, 107-126, 1951.

Biswas, S., C. E. Fichtel, and D. E. Guss, Study of the hydrogen, helium, and heavy nuclei in the Nov. 12, 1960, solar cosmic-ray event, Phys. Rev., 128, 2756-2771, 1962.

Biswas, S., C. E. Fichtel, D. E. Guss, and C. J. Waddington, Hydrogen, helium, and heavy nuclei from the solar event on Nov. 15, 1960, Goddard Space Flight Center, NASA, X-611-62-235, 1963.

Blamont, J. E., and C. deJager, Upper atmospheric turbulence near the 100 km level, Ann. Geophys., 17, 134-144, 1961.

Bourdeau, R. E., E. C. Whipple, Jr., J. L. Donley, and S. J. Bauer, Experimental evidence for the presence of helium ions based on Explorer 8 satellite data, J. Geophys. Res., 67, 467-475, 1962.

Brandt, J. C., and J. W. Chamberlain, Density of neutral gas in a planetary exosphere, Phys. Fluids, 3, 485-486, 1960.

Bryant, R., A comparison of theory and observation of the Echo 1 satellite, J. Geophys. Res., 66, 3066-3069, 1961.

Byram, E. T., T. A. Chubb, and H. Friedman, Dissociation of oxygen in the upper atmosphere, Phys. Rev., 98, 1594-1597, 1955.

Cahill, L. J., and P. G. Amazeen, The boundary of the geomagnetic field, J. Geophys. Res., 68, 1835-1843, 1963.

Carslaw, H., and J. C. Jaeger, Conduction of Heat in Solids, Oxford University Press, London, 1959.

Chapman, S., and T. G. Cowling, The Mathematical Theory of Non-Uniform Gases, Cambridge at the University Press, London, 1939.

Craig, H., and D. Lal, The production rate of natural tritium, Tellus, 13, 85-105, 1961.

Dalgarno, A., and D. Parkinson, Photoionization of atomic oxygen and atomic nitrogen, J. Atmospheric Terrest. Phys., 18, 335-337, 1960.

Damon, P. E., and J. L. Kulp, Inert gases and the evolution of the atmosphere, Geochim. Cosmochim. Acta, 13, 280-292, 1958.

Dessler, A. J., Ionospheric heating by hydromagnetic waves, J. Geophys. Res., 64, 397-401, 1959.

Dessler, A. J., and J. A. Fejer, Interpretation of K_p index and M-region geomagnetic storms, Planetary Space Sci., in press, 1963.

Dostrovsky, I., Z. Fraenkel, and L. Winsberg, Monte Carlo calculations of nuclear evaporation processes, 4, Spectra of neutrons and charged particles from nuclear reactions, Phys. Rev., 118, 781-793, 1960.

Dungey, J. W., The steady state of the Chapman-Ferraro problem in two dimensions, J. Geophys. Res., 66, 1043-1047, 1961.

Dungey, J. W., The interplanetary field and auroral theory, J. Phys. Soc. Japan, 17, Suppl. A-2, 15-18, 1962.

Eckart, C., Hydrodynamics of Oceans and Atmospheres, Pergamon Press, New York, 1960.

Fireman, E. L., Measurement of the (N, H^3) cross section in nitrogen and its relationship to the tritium production in the atmosphere, Phys. Rev., 91, 922-926, 1953.

Fireman, E. L., J. de Felice, and D. Tilles, Solar flare tritium in a recovered satellite, Phys. Rev., 123, 1935-1938, 1961.

Flamm, E., R. E. Lingenfelter, G. J. F. MacDonald, and W. F. Libby, Tritium and helium-3 solar flares and loss of helium from the earth's atmosphere, Science, 138, 48-59, 1962.

Francis, W. E., and R. Karplus, Hydromagnetic waves in the ionosphere, J. Geophys. Res., 65, 3593-3600, 1960.

Glueckauf, E., The composition of atmospheric air, Compendium of Meteorology, pp. 3-10, American Meteorological Society, Boston, 1951.

Goodman, N. R., On the joint estimation of the spectra, cospectrum, and quadrature spectrum of a two-dimensional stationary Gaussian process, Sci. Paper 10, Eng. Stabsl. Lab., New York University, 1957.

Greenhow, J. S., Eddy diffusion and its effect on meteor trails, J. Geophys. Res., 64, 2208-2209, 1959.

Grew, K. E., and T. L. Ibbs, Thermal Diffusion in Gases, Cambridge University Press, London, 1952.

Groves, G. V., Correlation of upper atmosphere air density with geomagnetic activity, Nov. 1960, Space Research, Proc. Intern. Space Sci. Symp., 2nd, Florence, 1961, pp. 751-753, North-Holland Publishing Company, Amsterdam, 1961.

Hale, L. C., Ionospheric measurements with a multigrid potential analyzer (abstract), J. Geophys. Res., 66, 1554, 1961.

Hanson, W. B., Upper-atmosphere helium ions, J. Geophys. Res., 67, 183-188, 1962.

Hanson, W. B., and F. S. Johnson, Electron temperatures in the ionosphere, Mem. Soc. Roy. Sci. Liege, 4, 390-424, 1961.

Harris, I., and W. Priester, Time-dependent structure of the upper atmosphere, J. Atmospheric Sci., 19, 286-301, 1962a.

Harris, I., and W. Priester, Theoretical models for the solar-cycle variation of the upper atmosphere, J. Geophys. Res., 67, 4585-4591, 1962b.

Herring, J., and L. Kyle, Density in a planetary exosphere, J. Geophys. Res., 66, 1980-1982, 1961.

Hide, R., and P. H. Roberts, The origin of the main geomagnetic field, Physics and Chemistry of the Earth, 4, edited by L. H. Ahrens, Frank Press, Kalervo Rankama, and S. K. Runcorn, pp. 27-98, Pergamon Press, London, 1961.

Hinteregger, H. E., and K. Watanabe, Photoionization rates in the E and F regions, 2, J. Geophys. Res., 67, 3373-3392, 1962.

Holland, H. D., Model for the evolution of the earth's atmosphere, Petrologic Studies: A Volume in Honor of A. F. Buddington, edited by A. E. J. Engel, Harold L. James, and B. F. Leonard, The Geological Society of America, 447-477, 1962.

Hurley, J., Interaction of a streaming plasma with the magnetic field of a two-dimensional dipole, Phys. Fluids, 4, 854-859, 1961.

Jacchia, L. G., The erratic orbital acceleration of 1957β, Sky and Telescope, 17, 278, 1958.

Jacchia, L. G., Recent advances in aeronomy: Reduction of orbital data of the satellites to yield atmospheric densities at heights up to 700 km, Nature, 183, 327, 1959.

Jacchia, L. G., A variable atmospheric-density model from satellite accelerations, Smithsonian Astrophys. Obs. Spec. Rept. 39, 15 pp., 1960.

Jacchia, L. G., A working model of the upper atmosphere, Nature, 192, 1147-1148, 1961.

Jacchia, L. G., and J. Slowey, Preliminary analysis of the atmospheric drag of the twelve-foot balloon satellite ($1961\delta_I$), Smithsonian Inst. Astrophys. Obs., Res. Space Sci., Spec. Rept. 84, 1962.

Jeans, J. H., Dynamical Theory of Gases, 4th edition, Cambridge University Press, London, 1925.

Karplus, R., W. E. Francis, and A. J. Dragt, The attenuation of hydromagnetic waves in the ionosphere, Planetary Space Sci., 9, 771-786, 1962.

King-Hele, D. G., and J. M. Rees, Scale height in the upper atmosphere, Proc. Roy. Soc. London, 270, 562-585, 1962.

King-Hele, D. G., and D. M. C. Walker, Density of the upper atmosphere and its dependence on the sun as revealed by satellite orbits, Nature, 186, 928-931, 1960.

King-Hele, D. G., and D. M. C. Walker, Upper-atmosphere density during the years 1957-1961 determined from satellite orbits, Space Research, Proc. Intern. Space Sci. Symp., 2nd, Florence, 1961, pp. 918-957, North-Holland Publishing Company, Amsterdam, 1961.

Kupperian, J. E., E. T. Byram, H. Friedman, and A. Unzicker, Molecular oxygen densities in the mesosphere over Fort Churchill, Ann. Intern. Geophys. Yr., 12, 440-444, 1960.

Lennard-Jones, E., Free paths in a non-uniform rarefied gas with an application to the escape of molecules from an isothermal atmosphere, Trans. Cambridge Phil. Soc., 22, 535-556, 1923.

MacDonald, G. J. F., Calculations on the thermal history of the earth, J. Geophys. Res., 64, 1967-2000, 1959.

MacDonald, G. J. F., Spectrum of hydromagnetic waves in the exosphere, J. Geophys. Res., 66, 3639-3670, 1961.

MacDonald, G. J. F., On the internal constitution of the inner planets, J. Geophys. Res., 67, 2945-2974, 1962.

McDonald, Frank, Solar Proton Manual, Goddard Space Flight Center, NASA, X-611-62-122, 1963.

Munk, W. H., and G. J. F. MacDonald, Rotation of the Earth, A Geophysical Discussion, Cambridge University Press, London, 1960.

Ness, N. F., T. L. Skillman, C. S. Scearce, and J. P. Heppner, Magnetic field fluctuations on the earth and in space, International Conference on Cosmic Rays and the Earth Storm, part 2, J. Phys. Soc. Japan, 17, Suppl. A-2, 27-33, 1962.

Neugebauer, M., and C. W. Snyder, Solar plasma experiment, Science, 138, 1095-1096, 1962.

Nicolet, M., The aeronomic problem of helium, Ann. Geophys., 13, 1-21, 1957.

Nicolet, M., Constitution of the atmosphere at ionospheric levels, J. Geophys. Res., 64, 2092-2101, 1959.

Nicolet, M., Les variations de la densité et du transport de chaleur par conduction dans l'atmosphère supérieure, Space Research, Proc. Intern. Space Sci. Symp., 1st, Nice, 1960, edited by H. K. Kallmann-Bijl, pp. 46-89, North-Holland Publishing Company, Amsterdam, 1960.

Nicolet, M., Structure of the thermosphere, Planetary Space Sci., 5, 1-32, 1961a.

Nicolet, M., Les modèles atmosphériques et l'hélium, Space Research, Proc. Intern. Space Sci. Symp., 2nd, Florence, 1961, pp. 896-901, North-Holland Publishing Company, Amsterdam, 1961b.

Nicolet, M., Helium, an important constituent in the lower exosphere, J. Geophys. Res., 66, 2263-2264, 1961c.

Opik, E. J., and S. F. Singer, Distribution of density in a planetary exosphere, Phys. Fluids, 2, 653-655, 1959.

Opik, E. J., and S. F. Singer, Distribution of density in a planetary exosphere, Phys. Fluids, 3, 486-488, 1960.

Paetzold, H. K., A preliminary model for the variations of upper air densities, Nature, 190, 35-36, 1961.

Polkhunkov, A., Mass-spectrometer investigations of the structural parameter of the earth's atmosphere at altitudes from 100 to 210 km, Planetary Space Sci., 9, 269-279, 1962.

Priester, W., Sonnenaktivitaet und Abbremsung der Erdsatelliten, Mitt. Univ. Sternwarte Bonn, No. 24, 4 pp., and Naturwissenschaften, 46, 197-198, 1959.

Priester, W., Solar activity effect and diurnal variation in the upper atmosphere, J. Geophys. Res., 66, 4143-4148, 1961.

Priester, W., and H. A. Martin, Solare und tageszeitliche effekte in der hochatmosphaere aus beobachtungen kunstlicher erdsatelliten, Mitt. Univ. Bonn. No. 29, 53 pp., English translation: Roy. Aircraft Estab., Farnborough Library, Transl. 901, 20 pp., 1960.

Priester, W., H. A. Martin, and K. Kramp, Diurnal and seasonal density variations in the upper atmosphere, Nature, 188, 202-204, 1960.

Rossi, B., The interplanetary plasma, paper presented at American Physical Society Meeting, New York, 1962.

Rubey, W. W., Geologic history of sea water: An attempt to state the problem, Bull. Geol. Am., 62, 1111-1148, 1951.

Rubey, W. W., Development of the hydrosphere and atmosphere, with special reference to probable composition of the early atmosphere, in Crust of the earth, edited by Arie Poldevaart, Geol. Soc. Am. Spec. Paper 62, 631-650, 1955.

Schaeffer, E., The dissociation of oxygen measured by a rocket-borne mass spectrometer, J. Geophys. Res., 68, 1175-1176, 1963.

Schaeffer, O. A., and J. Zahringer, Solar flares: Helium in satellite materials, Phys.Rev. Letters, 8, 389-390, 1962.

Sharp, G. W., W. B. Hanson, and D. D. McKibbin, Atmospheric density measurements with a satellite-borne microphone gage, J. Geophys. Res., 67, 1375-1382, 1962.

Slutz, R. J., The shape of the geomagnetic field boundary under uniform external pressure, J. Geophys. Res., 67, 505-513, 1962.

Smith, E. J., Leverett Davis, Jr., P. J. Coleman, Jr., and C. P. Sonett, Magnetic field, Science, 139, 909, 1963.

Spitzer, L., The terrestrial atmosphere above 300 km, The Atmospheres of the Earth and Planets, edited by G. P. Kuiper, pp. 211-247, University of Chicago Press, Chicago, 1949.

Sugiura, M., Evidence of low-frequency hydromagnetic waves in the exosphere, J. Geophys. Res., 66, 4087-4095, 1961.

Townsend, J. W., Jr., and E. B. Meadows, Density of the winter nighttime Arctic upper atmosphere, Ann. Geophys., 14, 117-130, 1958.

Townsend, J. W., Jr., E. B. Meadows, and E. C. Pressly, A mass spectrometric study of the upper atmosphere, in Rocket Exploration of the Upper Atmosphere, Pergamon Press, London, 169-188, 1954.

Turekian, K. K., The terrestrial economy of helium and argon, Geochim. Cosmochim. Acta, 17, 37-42, 1959.

Urey, H. C., The atmospheres of the planets, Encyclopedia of Physics, Vol. 52, edited by S. Flugge and Marburg, pp. 363-418, Springer-Verlag, Berlin, 1959.

Wasserburg, G. J., E. Mazor, R. E. Zartman, Isotopic and chemical composition of some terrestrial natural gases, Earth Science and Meteoritics, compiled by J. Geiss and E. D. Goldberg, pp. 219-240, North-Holland Publishing Company, Amsterdam, 1963.

Wilson, A. T., and G. J. Fergusson, Origin of terrestrial tritium, Geochim. Cosmochim. Acta, 18, 273-277, 1960.

Zadunaisky, P. E., I. I. Shapiro, and H. M. Jones, Experimental and theoretical results on the orbit of Echo 1, Smithsonian Astrophys. Obs. Spec. Rept. 61, 22 pp., 1961.

Zartman, R. E., G. J. Wasserburg, and J. H. Reynolds, Helium, argon, and carbon in some natural gases, J. Geophys. Res., 66, 277-306, 1961.

HYDROMAGNETIC WAVES IN SPACE

by

Gordon J. F. MacDonald

Institute of Geophysics
and Planetary Physics
University of California
Los Angeles, California

LIST OF FIGURES

	Caption	After Page
Figure 1:	Diagnostic diagram for hydromagnetic wave propagation in a uniform fluid.	29
Figure 2:	Slowness surfaces for propagation of hydromagnetic waves in a uniform medium.	29
Figure 3:	Surfaces of constant phase for a uniform medium.	30
Figure 4:	Variation of free period of oscillation with geomagnetic latitude.	39

HYDROMAGNETIC WAVES IN SPACE

Gordon J. F. MacDonald
Institute of Geophysics
and Planetary Physics
University of California
Los Angeles, California

Since the work of Herlofson (1950) and van de Hulst (1951) it has been known that in a compressible, non-dissipative, hydromagnetic medium, small amplitude plane waves are propagated at three distinct speeds corresponding to three different polarizations of the fluid displacement. The theory of hydromagnetic waves has been discussed by many authors. Lighthill (1960) developed a comprehensive theory for the propagation of small amplitude waves in a spatially homogeneous medium. The propagation of weak hydromagnetic discontinuities has been considered in an elegant fashion by Bazer and Fleischman (1959a,b) and in particular these authors treat propagation in an inhomogeneous medium. Bazer and Hurley (1963) developed a complete theory of geometrical hydromagnetics and their results are of direct applicability to the study of the propagation of relatively high-frequency hydromagnetic waves in the inhomogeneous ionized gas surrounding the earth.

It is readily shown that the wave transmitted at a speed intermediate to that of the other two is purely transverse in that both the fluid displacement and the magnetic disturbance are normal to the direction of propagation and to the initial undisturbed magnetic field. This non-compressive, pure shear wave is commonly called the intermediate or Alfvén wave.

The other two types of propagating disturbances involve displacements that are co-planar with the propagation vector and the undisturbed field. These waves involve compression which alters the thermodynamic state of the medium and they are labelled magneto-acoustic waves.

While the theoretical development of the propagation of hydromagnetic waves has been rapid, there has been relatively less work concerned with the experimental study of hydromagnetic wave propagation. The laboratory observation of hydromagnetic waves requires an ionized medium with a magnetic field of such strength that the magnetic restoring forces are comparable to the ordinary elasticity of the fluid. These conditions are difficult to meet. The earth, however, is surrounded by a low-density, partially ionized medium in which the earth's magnetic field constrains the particle motion. This region (magneto-sphere) gives way to interplanetary space dominated by streaming plasma issuing from the sun. In both regions, the magnetic restoring forces are significant and true hydromagnetic waves can propagate.

It has long been known that the magnetic field at the earth's surface undergoes a wide variety of disturbances. Many of these disturbances have the character of an almost periodic variation with periods ranging from tenths of seconds to several hundred seconds. These oscillations have been interpreted as due to the propagation of magnetohydrodynamic waves in the

magnetosphere through the ionosphere and into the neutral atmosphere. At the boundary of the ionosphere, the medium changes from a partially ionized gas to a wholly neutral gas. Part of the energy contained within the traveling hydromagnetic wave is transmitted as an electromagnetic wave. At the relevant low frequencies, the wavelength of the generated electromagnetic wave is very large.

In addition to ground observations, many measurements have been made of the magnetic field in nearby space, both within and outside the magnetosphere. These measurements again demonstrate the existence of almost periodic variations of the magnetic field, and attempts have been made to interpret these observations in terms of the propagating or standing hydromagnetic waves.

In a plasma, a large variety of waves can be propagated. We will be concerned solely with those waves which are truly magnetohydrodynamic. Thus, we restrict our consideration to waves having frequency and wavelength which are sufficiently large so that the plasma acts as a continuum. In the following section, we describe conditions that exist within the outer regions of the earth's atmosphere and in neighboring space so as to delineate the frequency-wave number space for which the continuum approximations are appropriate. In the following section, we take up the description of waves in a compressible, non-dissipative, hydromagnetic medium. The detailed geometry

of the wave solutions are described and the observable properties of the waves are considered in some detail. The problem of the refraction and reflection of hydromagnetic waves at sharp boundaries is considered from an elementary point of view. The questions of propagation in an inhomogeneous medium are not taken up in detail since they have been treated by Bazer and Hurley (1963). Furthermore, the influence of the earth's gravity field is neglected as this has been considered by MacDonald (1961).

The observations of hydromagnetic waves at the surface of the earth and in satellites are next considered. The observations are by no means complete and are unsatisfactory in many features. The spectral characteristics of the vector field remain, for the most part, unknown. Moreover, the lack of uniform instrumentation and the dearth of digital recording magnetometers at the earth's surface limits the description of high-frequency magnetic variations at the earth's surface. There are similar difficulties with the description of hydromagnetic waves in space. The interaction of the solar wind with the earth's magnetic field undoubtedly gives rise to much of the magnetic activity observed at the earth's surface. In addition, turbulent motion of a plasma can generate a magnetic field and, in the process, radiate energy in the form of magnetohydrodynamic waves. However, there is no adequate theory for the source of the observed magnetic fluctuation.

Plasma in the Magnetosphere and Nearby Space

A summary of the conditions existing in the outer magnetosphere and in interplanetary space as derived by these observations is shown in Table I along with average conditions in the solar corona. In addition, the important parameters describing the plasma are also listed.

The region between 100 km and the boundary of the earth's magnetic field is filled with a partially ionized gas of low density. In the lower part of the region, the concentration of the ionized particles is small compared with the neutrals, but at higher altitudes, the balance shifts with ionized particles dominating. A typical value for the concentration of ionized particles at 1000 km is 10^4 O^+ with somewhat fewer He^+ and less than 100 H^+ all at a temperature of 1500°K. At 25,000 km, the ionized gas would consist of about 100 H^+ cm^{-3}. The temperature and abundance of neutral and ionized particles changes with a diurnal cycle and with the solar cycle. In the outer regions, the conditions are most uncertain since they have been studied only through the propagation of whistlers (see also, Gringauz, 1962a). In addition, calculations of the vertical distribution, assuming a diffusive equilibrium, have been carried out (Nicolet, 1963).

The conditions in nearby space have been probed by a number of Soviet vehicles (Gringauz, 1962b) and by Explorer 10 and Mariner 2. The most detailed results have been published by

TABLE I

Parameters Describing Plasmas

	Outer Magnetosphere	Nearby Space Lunik 2, Lunik 3, Explorer 10	Interplanetary Space Mariner 2	Solar Corona
Density (cm^{-3})	50 - 500	10	2 - 5	10^6
Magnetic Field (gauss)	2×10^{-4} - 2×10^{-3}	1.5×10^{-4}	5×10^{-5}	10^{-4}
Temperature (°K)	1000 - 2000	10^5	2 - 8 $\times 10^5$	10^6
Plasma Frequency (sec^{-1})	4.0×10^5 - 1.3×10^6	1.8×10^5	8.0×10^4 - 2.2×10^5	5.6×10^7
Debye Length (cm)	10 - 43	6.9×10^2	1.3×10^3 - 4.3×10^3	10^{-4}
Particles in Debye Sphere (cm)	2.1×10^6 - 4.0×10^6	1.4×10^{10}	4.6×10^{10} - 1.6×10^{11}	4×10^2
Collision Frequency (sec^{-1})	0.3 - 10	2×10^{-4}	2.6×10^{-6} - 5.9×10^{-5}	5×10^{-2}
Mean Free Path (cm)	1.8×10^6 - 6.3×10^7	6×10^{11}	1×10^{13} - 1.6×10^{14}	6×10^8
Electron Gyro Frequency (sec^{-1})	3.5×10^3 - 3.5×10^4	2.6×10^3	8.8×10^2	1.6×10^3
Proton Gyro Frequency (sec^{-1})	1.9 - 19	1.4	4.8×10^{-1}	0.9
Larmor Radius Electron (cm)	5×10^2 - 7×10^3	6.6×10^4	2.8×10^5 - 5.6×10^5	3.7×10^5
Larmor Radius Protons (cm)	2×10^4 - 3×10^5	2.8×10^6	1.2×10^7 - 2.4×10^7	1.6×10^7

Bonetti et al. (1963) on the plasma observations and by Hepner et al. (1963) on the magnetic observations made by instruments aboard Explorer 10. Preliminary reports have been given by Neugebauer and Snyder (1963) and by Coleman et al. (1962) on the plasma and magnetic field observations made by Mariner 2. The results obtained by these vehicles are in general agreement with those obtained by the Soviet observations (Bonetti et al., 1963).

In a plasma, the Debye length is defined as the distance over which the removal of all charged particles of one sign would result in an electric field comparable to kT. In the outer magnetosphere, the Debye length is a few tens of centimeters while in nearby space, it is on the order of tens of meters. These distances are large compared with the Debye length in the solar corona which is on the order of 10^{-4} cm. From the particle density and the Debye length, the number of particles in a Debye sphere can be calculated. In the magnetosphere, the number of particles is on the order of 10^{6} so that a charged particle can interact with another charged particle at distances some 10^{6} particle diameters away. The number of particles in a Debye sphere in interplanetary space is even greater, on the order of 10^{10}. The product of the Debye length and the interparticle spacing is proportional to the square root of the ratio of the mean kinetic and interparticle potential energies. When the mean interparticle potential energies are much greater than kinetic energies, the ions and electrons can

recombine. For plasmas in the interplanetary space and in the upper magnetosphere, the Debye length is far greater than the mean interparticle spacing concentration. Thus, in these plasmas there is little tendency to recombine and the gases are highly ionized.

A quantity of interest is the Debye frequency. We consider a steady initial state in which there is an initial number density of n electrons cm^{-3} neutralized by an equilibrium number of ions. The electron density is disturbed and we are interested in the subsequent behavior of the disturbance in density. The disturbance will oscillate with an angular frequency ω_p, given by $\omega_p^2 = 4\pi ne^2/m$. The plasma frequency in the outer magnetosphere is on the order of 10^6 c/s. The plasma frequency in nearby space is somewhat less, on the order of 1×10^5 c/s. The plasma frequency is not of direct interest in studies of the propagation of hydromagnetic waves, but since it determines the free oscillations of the plasma, it is clear that only frequencies much greater or much less then the natural response frequency of the plasma can be propagated.

While the plasma in nearby space is almost certainly nearly completely ionized, the plasma in the upper ionosphere and lower magnetosphere is partially ionized. In a partially ionized gas, there are several mean free paths corresponding to neutral-neutral collisions, neutral-charged particle collisions, etc. Inelastic processes, such as ionization and recombination

introduce further complications. The neutral-neutral mean free path is the distance a neutral particle must travel before it exchanges its momentum by a direct collision with another neutral particle. The difference between the neutral-neutral mean free path and the neutral-charged particle mean free path is small. Over the entire magnetosphere, the neutral gas mean free path extends several kilometers and is large compared with other lengths that characterize the motion of the charged particles.

As a result of Coulomb interactions, a charged particle exchanges its momentum, not by single collisions, but in small bits through distant interactions. The large number of particles within a Debye radius of a charged particle indicates the number of possible interactions. Because of the small but frequent exchanges of momentum, a particle may transfer an appreciable portion of its momentum in distances that are short compared to the distance a particle must travel to exchange all its momentum. In the outer magnetosphere, the mean free paths are on the order of tens to hundreds of kilometers. In space, the mean free paths range from substantial fractions to tens of astronomical units. Thus, in nearby space, the mean free path is long and the collision frequency between particles is very low.

In a magnetic field, a second length becomes important, the Larmor radius, r_L, or the radius of the circle formed by the particle orbit about the magnetic lines of force, $r_L = nvc/eB$,

where v is the velocity tangential to the circular orbit and B is the magnetic field. For a proton, the ratio of the Larmor radius to the mean free path is a very small number over the outer magnetosphere and nearby space. Under these conditions, the behavior of the plasma is determined by the interaction of the charged particles with the magnetic field rather than the interaction between pairs of particles. In the magnetosphere, the collisions can be of importance, but in nearby space, the plasma acts very nearly as a collisionless medium. The hydromagnetic behavior thus must be sought in the interaction between the independent charged particles and the macroscopic electric and magnetic fields.

Table II lists the characteristic velocities for the various regions. It is seen that over the regions of interest, the Alfvén wave velocity is much larger than the sound wave velocity. The ratio of the sound wave to the Alfvén wave velocity, β, is a small number on the order of 10^{-1} to 10^{-3}. The ratio of charged particle pressure to magnetic pressure is of the order of β^2. Thus, over the magnetosphere and in nearby space, magnetic pressure dominates the particle pressure. In the streaming plasma of interplanetary space, the dynamic pressure ρv^2 is much larger than the thermodynamic pressure.

The usual conditions for the validity of the continuum hydromagnetic approximation are obtained by deriving the single-fluid equations with a scalar conductivity from the two-fluid

TABLE II

Characteristic Velocities in Plasma

	Outer Magnetosphere	Nearby Space Lunik 2, Lunik 3, Explorer 10	Interplanetary Space Mariner 2	Solar Corona
Thermal Velocities ($cm\ sec^{-1}$)				
Electron	$1.7 \times 10^7 - 2.5 \times 10^7$	1.7×10^8	$2.5 \times 10^8 - 4.9 \times 10^8$	5.3×10^8
Proton	$4.0 \times 10^5 - 5.7 \times 10^5$	4.0×10^6	$5.7 \times 10^6 - 1.1 \times 10^7$	1.2×10^7
Sound Velocity ($cm\ sec^{-1}$)	$3.6 \times 10^5 - 5.1 \times 10^5$	3.6×10^6	$5.1 \times 10^6 - 1.0 \times 10^7$	1.1×10^7
Alfvén Velocity ($cm\ sec^{-1}$)	$6.1 \times 10^6 - 2.0 \times 10^7$	1.0×10^7	$4.8 \times 10^6 - 2.7 \times 10^6$	2.2×10^4

theory. The restrictions of the continuum theory are that the impressed frequency of the wave must be much less than the electron gyro-frequency, the frequency of collision between the two species of particles, and the plasma frequency. In addition, it is required that the electron gyro-frequency must be much less than the collision frequency. These requirements are in addition to the usual requirement that the displacement current be negligible im comparison with the conduction currents. As can be seen from the numbers listed in Table I, these conditions are not met either within the outer magnetosphere or nearby space. The electron gyro-frequency is much greater than the collision frequency. The condition that the impressed frequency be less than the collision frequency holds only for very low-frequency waves. However, these conditions apply to the general treatment of hydromagnetic waves and it is possible that certain waves can propagate under conditions holding within the magnetosphere and nearby space. In particular, the intermediate or purely transverse waves involve an interaction of the particles through the intervention of the magnetic field. There is no thermodynamic change in the medium and the absence of particle collisions is of no consequence. Thus, this mode can be treated by the continuum approximation even when the collision frequency is far less than the gyro-frequencies.

Description of Hydromagnetic Waves

We consider a fluid in which there exists a scalar fluid pressure, P', which is a function of the entropy, S', and the

density, ρ'. The displacement current is neglected and Ohm's law is taken in the form $\underset{\sim}{J}' = \underset{\sim}{B}' \times \underset{\sim}{V}'$ where $\underset{\sim}{J}'$ is the current, $\underset{\sim}{B}'$ is the magnetic field, and $\underset{\sim}{V}'$ is the velocity. Dissipation due to the finite viscosity and thermal conductivity of the fluid is neglected. The action of additional external force fields are also neglected. Under these conditions, the equation of motion, Maxwell's equations, the continuity equation, and the conservation of energy equation become

$$\frac{d\underset{\sim}{V}'}{dt} = -\frac{1}{\rho'}\nabla P' - \frac{1}{4\pi\mu\rho'}\underset{\sim}{B}' \times (\nabla \times B') \qquad (1)$$

$$\nabla \cdot \underset{\sim}{B}' = 0$$

$$\frac{d\underset{\sim}{B}'}{dt} = \nabla \times (\underset{\sim}{V}' \times \underset{\sim}{B}')$$

$$\frac{d\rho'}{dt} + \rho' \nabla \cdot V' = 0$$

$$\frac{dS'}{dt} = \frac{Q'}{T'} \quad ,$$

where Q' is the net accession of heat and provides a source term.

The equation of the conservation of mass can be combined with thermodynamic relations to relate pressure and density variations

$$\frac{dP'}{dt} - C'^2 \frac{d\rho'}{dt} = \rho \frac{(\gamma'-1)}{\alpha'} \frac{Q'}{T'} \quad , \qquad (2)$$

where C' is the velocity of sound while $\gamma' = c'_p/c'_v$ is the ratio of the specific heats.

The set of nonlinear equations given in (1) can be reduced to a tractable linear form by the perturbation scheme

$$\begin{aligned} \underset{\sim}{V}' &= \underset{\sim}{V} + \underset{\sim}{v} + \ldots \\ \underset{\sim}{B}' &= \underset{\sim}{B} + \underset{\sim}{b} + \ldots \\ P' &= P + p + \ldots \\ S' &= S + s + \ldots \\ \rho' &= \rho_o + \rho + \ldots \\ Q' &= q + \ldots \end{aligned} \qquad (3)$$

The zeroth order state of the medium is characterized by a velocity $\underset{\sim}{V}$, a magnetic field $\underset{\sim}{B}$, a pressure P, an entropy S, and a density ρ_o. Thus, in the absence of a disturbance, it is assumed that the medium is undergoing steady isotropic flow and the equations governing this flow are

$$\rho_o (\underset{\sim}{V}\cdot\nabla)\,\underset{\sim}{V} + C^2\nabla\rho_o + \frac{1}{4\pi\mu}\,\underset{\sim}{B}\times(\nabla\times\underset{\sim}{B}) = 0 \qquad (4)$$

$$\begin{aligned} &\nabla\cdot\underset{\sim}{B} = 0 \\ &\nabla\times(\underset{\sim}{V}\times\underset{\sim}{B}) = 0 \\ &\nabla\cdot(\rho\underset{\sim}{V}) = 0 \\ &\nabla\cdot S = 0 \end{aligned} \quad .$$

We note that if in the initial state the velocity field $\underset{\sim}{V}$ vanishes, Maxwell's equations are satisfied identically. The set of equations (4) describes the time-independent zeroth-order state upon which a disturbance is imposed.

Substituting (3) into (1), we obtain the time-dependent linearized equation describing the propagation of the disturbance

$$\rho_o \frac{\partial \underset{\sim}{v}}{\partial t} + \rho_o (\underset{\sim}{V}\cdot\nabla)\,\underset{\sim}{v} + \rho_o (\underset{\sim}{v}\cdot\nabla)\,\underset{\sim}{V} + \rho\,\underset{\sim}{V}\cdot\nabla\,\underset{\sim}{v} \qquad (5)$$

$$= -\frac{1}{4\pi\mu}\,\underset{\sim}{B}\times(\nabla\times\underset{\sim}{b}) - \frac{1}{4\pi\mu}\,\underset{\sim}{b}\times(\nabla\times\underset{\sim}{B}) - c^2\,\nabla\rho$$

$$\nabla\cdot\underset{\sim}{b} = 0$$

$$\frac{\partial \underset{\sim}{b}}{\partial t} = \nabla\times(\underset{\sim}{V}\times\underset{\sim}{b}) + \nabla\times(\underset{\sim}{v}\times\underset{\sim}{B})$$

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\underset{\sim}{V}\rho) + \nabla(\rho_o\underset{\sim}{v}) = 0$$

$$\frac{ds}{dt} + \underset{\sim}{v}\cdot\nabla S = \frac{q}{T} \quad .$$

For many purposes, the plasma within the magnetosphere can be regarded as stationary and, therefore, the time independent velocity $\underset{\sim}{V}$ vanishes. This is not true of the plasma issuing from the sun and streaming by the earth. In the magnetosphere, the magnetic field varies as does the density of the ionized medium due to the combined effects of gravity and varying magnetic field. The propagation of disturbances at high

frequencies in an inhomogeneous medium have been considered in detail by Bazer and Hurley (1963). The question of the propagation of a disturbance in a moving medium has been treated by Fejer (1963). For the present purposes, we wish to consider only disturbances propagated in a stationary, homogeneous medium.

Equations (5) reduce to

$$\rho_o \frac{\partial \underset{\sim}{v}}{\partial t} = - \frac{1}{4\pi\mu} \underset{\sim}{B} \times (\nabla \times \underset{\sim}{b}) - c^2 \nabla \rho \qquad (6)$$

$$\nabla \cdot \underset{\sim}{b} = 0$$

$$\frac{\partial \underset{\sim}{b}}{\partial t} = \nabla \times (\underset{\sim}{v} \times \underset{\sim}{B})$$

$$\frac{\partial \rho}{\partial t} + \rho_o \nabla \cdot \underset{\sim}{v} = 0$$

$$\frac{\partial s}{\partial t} = \frac{q}{T} \quad ,$$

when the medium is stationary and the steady magnetic field is constant in space. This set of equations has been studied extensively and the general problems of propagation and disturbances governed by these equations are relatively well understood.

The energy equation can be obtained by multiplying the equation of motion by $\underset{\sim}{v}$, Maxwell's equations by $\underset{\sim}{b}$, and the

equation of mass conservation by $C^2\rho/\rho_o$. Adding the results, we obtain the energy equation in the form

$$\frac{\partial E}{\partial t} + \nabla\cdot\underset{\sim}{F} = 0 \quad , \tag{7}$$

where E is the energy and $\underset{\sim}{F}$ is the flux

$$E = \frac{1}{2}\rho_o v^2 + \frac{1}{8\pi\mu} b^2 + \frac{1}{2}\frac{C^2}{\rho_o}\rho^2 \tag{8}$$

$$\underset{\sim}{F} = C^2\rho\underset{\sim}{v} + \frac{1}{4\pi\mu}(B \times \underset{\sim}{v}) \times \underset{\sim}{b} \quad .$$

The last term in the expression for the flux is the familiar Poynting vector, since in a fluid of infinite conductivity, the electric field is given by $\underset{\sim}{B} \times \underset{\sim}{v}$. The first term in the flux is the ordinary flux term of acoustics.

A principal difficulty in the study of (6) is concerned with the anisotropy introduced by the magnetic field. The magnetic field restricts the motion of individual particles perpendicular to the field, but provides no restriction in a direction parallel to the field. As a result, the propagation is direction-dependent and the phase and group velocities of the waves depend on the relative orientation of the propagation vectors and the magnetic field. The description of propagation in an isotropic medium is aided by the construction of slowness surfaces in wave number space at a fixed frequency and the inversion of these surfaces to form surfaces of constant phase.

The propagation of a wave can be described in terms of a phase function $\varphi(\underset{\sim}{x}, t)$. The surface

$$\varphi(\underset{\sim}{x}, t) = \text{constant}$$

at a given time might correspond to a wavecrest or trough. The phase velocity $\underset{\sim}{C}$ is the velocity of the surface, the wave number $\underset{\sim}{k}$ is defined as the normal to the phase surface

$$\underset{\sim}{k} = \nabla \varphi \quad ,$$

and is the direction of the phase velocity. The magnitude of the phase velocity is

$$\underset{\sim}{C} = - \frac{\partial \varphi}{\partial t} \Big/ \nabla \varphi = \frac{\omega}{\underset{\sim}{k}} \quad ,$$

where the frequency ω is defined by $\omega = \partial\varphi/\partial t$. The wavecrests and troughs are at right angles to the vector $\underset{\sim}{k}$ and the wavelength is $2\pi/|\underset{\sim}{k}|$.

The equations for a given problem define a dispersion relation between $\underset{\sim}{k}$, $\underset{\sim}{x}$, and ω. For convenience, the dispersion relation is written in the form

$$W(\underset{\sim}{k}, \underset{\sim}{x}, \omega) = 0 \quad .$$

The group velocity

$$\underset{\sim}{U} = \frac{d\underset{\sim}{x}}{dt} \quad ,$$

is determined from

$$\underset{\sim}{U} = - \frac{\partial W}{\partial \underset{\sim}{k}} \Big/ \frac{\partial W}{\partial \omega} \quad .$$

The existence of a scalar function φ corresponds to the conservation of the number of waves, or, equivalently, of the wave number $\underset{\sim}{k}$. The existence of a function φ implies

$$\frac{\partial^2 \varphi}{\partial \underset{\sim}{x} \partial t} - \frac{\partial^2 \varphi}{\partial t \partial \underset{\sim}{x}} = 0 \quad ,$$

or

$$\frac{\partial \underset{\sim}{k}}{\partial t} + \nabla \omega = 0 \quad .$$

Introducing the group velocity, we show that the equation for the conservation of wave number is

$$\frac{\partial \underset{\sim}{k}}{\partial t} + (\underset{\sim}{U} \cdot \nabla) \underset{\sim}{k} = - \nabla W \Big/ \frac{\partial W}{\partial \omega} \quad ,$$

in which use has been made of the fact that $\underset{\sim}{k}$ is an irrotational vector. The relation for the equation of conservation of wave number can be regarded as a form of Hamilton's equation. On the lefthand side is the rate of change of wave number following a point moving with the group velocity $\underset{\sim}{U}$ which is the velocity of a group of waves having the wave number $\underset{\sim}{k}$. The number of

waves is conserved provided the dispersion relation is independent of the spatial coordinate; in this case, the righthand side vanishes. In general, the coefficients of the partial differential equation describing the general physical situation will be a function of position and the righthand side will be finite. The wave number and frequency must be then interpreted as the local propagation vector in frequency. These quantities can be obtained by local arguments and their propagation in space is then fixed by the above equation. Once $\underset{\sim}{k}$ is determined in space by the dispersion relations and its propagation by the equation for the conservation of wave number, it is possible to determine the phase function.

At a given frequency and point in space, it is possible to construct a three-dimensional surface and wave number space determined by

$$W(\omega, \underset{\sim}{k}, \underset{\sim}{x}) = 0 \quad .$$

This surface will be termed the slowness surface (Synge, 1957). Every point on the slowness surface corresponds to a wave. Some of these waves, however, are disallowed because of the radiation condition which allows only the propagation of energy outward from a source in an initial value problem.

The surface of constant phase which can be constructed from the slowness surface is the envelope at a given time of

waves originating from a source at a time t = 0. The direction of the wave resulting from a point on the slowness surface is along the normal to the slowness surface. It should be noted that the direction of the wave number $\underset{\sim}{k}$ and the normal to the slowness surface need not be the same. The crests and troughs of the waves are then not necessarily at right angles to the direction of propagation. Since the normal to the slowness surface is parallel to the direction $\underset{\sim}{x}$, we have

$$\underset{\sim}{x} = \text{constant } \frac{\partial W}{\partial \underset{\sim}{k}} \quad ,$$

where, at unit time,

$$\underset{\sim}{k} \cdot \underset{\sim}{x} = \omega \quad .$$

Combining the last two equations gives the surface of constant phase at unit time

$$\underset{\sim}{x} = - \frac{\partial W}{\partial \underset{\sim}{k}} \Big/ \frac{\partial W}{\partial \omega} \quad .$$

At unit time, the surface of constant phase is then a representation of the group velocity. This interpretation runs into difficulty, however, when the group velocity is a function of the spatial coordinate; in that case, only the initial wave form can be calculated and the wave form is distorted as the wave progresses through the medium. A further complication arises from dispersion. Waves are dispersed if $W(\underset{\sim}{k}, \underset{\sim}{x}, \omega)$

is not homogeneous in $\underset{\sim}{k}$ and ω; the phase surface for different frequencies will spread in time. Once the slowness surface is determined, the surface of constant phase can be calculated. A geometrical interpretation of the slowness surface can then be used to describe the motion of the wave.

The set of equations (6) can be studied by examining solutions which are proportional to

$$\exp\left[i\left(\underset{\sim}{k}\cdot\underset{\sim}{x} - \omega t\right)\right] \quad .$$

These solutions describe the propagation of plane waves with a wave vector $\underset{\sim}{k}$ and a frequency ω. The set of equations (6) then yield the algebraic relations

$$\rho_o\ \omega\underset{\sim}{v}^* = \frac{1}{4\pi\mu}\ \underset{\sim}{B} \times \left(\underset{\sim}{k} \times \underset{\sim}{b}^*\right) + c^2\ \underset{\sim}{k}\ \rho^* \qquad (9)$$

$$\omega\underset{\sim}{b}^* = \underset{\sim}{k} \times \left(\underset{\sim}{B} \times \underset{\sim}{v}^*\right)$$

$$\omega\rho^* = \rho_o\ \underset{\sim}{k}\cdot\underset{\sim}{v}^* \quad ,$$

where the * quantities refer to the complex amplitudes. The density in Equation (9) can be eliminated using the continuity equation and we obtain

$$\rho_o\ \omega\ \underset{\sim}{v}^* = \frac{1}{4\pi\mu}\ \underset{\sim}{B} \times \left(\underset{\sim}{k} \times \underset{\sim}{b}^*\right) + \frac{c^2\ \rho_o}{\omega}\ \left(\underset{\sim}{k}\cdot\underset{\sim}{v}^*\right)\ \underset{\sim}{k} \quad . \qquad (10)$$

We also introduce the phase velocity of the wave $c = \omega/|\underset{\sim}{k}|$. Equation (9) then takes the form

$$c\,\underset{\sim}{b}^* = \underset{\sim}{n} \times (\underset{\sim}{B} \times \underset{\sim}{v}^*) \tag{11}$$

$$\rho_o\, c\, \underset{\sim}{v}^* = \frac{1}{4\pi\mu}\, \underset{\sim}{B} \times (\underset{\sim}{n} \times \underset{\sim}{b}^*) + \frac{c^2}{c}\, \rho_o\, (\underset{\sim}{n}\cdot\underset{\sim}{v}^*)\, \underset{\sim}{n} \quad ,$$

where $\underset{\sim}{n}$ is a unit vector in the direction of propagation.

Equations (11) simplify considerably for special geometries. Consider the wave vector $\underset{\sim}{k}$ to be along the x_1 axis with the plane of the wave vector $\underset{\sim}{k}$ and magnetic field as the $\underset{\sim}{B}$ x_1 - x_2 plane. In this case, the equations reduce to

$$cb_2^* = v_1^* B_2 - v_2^* B_1 \tag{12}$$

$$cb_3^* = - v_3^* B_1$$

$$v_1^* \left(c - \frac{c^2}{c}\right) = \frac{B_2 b_2^*}{4\pi\mu\rho_o}$$

$$cv_2^* = - \frac{B_1 b_2^*}{4\pi\mu\rho_o}$$

$$cv_3^* = - \frac{B_1 b_3^*}{4\pi\mu\rho_o} \quad .$$

The equations separate into two groups, one involving only h_3^* and v_3^* and a second group involving the velocity and magnetic

disturbances in the x_1 - x_2 plane. The perturbations of the two groups of variables are thus propagated independently.

In the first group, the component h_3 in the magnetic field which is perpendicular to the direction of propagation and to the constant field $\underset{\sim}{B}$ oscillates and at the same time the velocity v_3 varies. The phase velocity is

$$c_i = B_1 / \sqrt{4\pi\mu\rho_o} \quad , \tag{13}$$

which, in general, requires that the frequency ω be related to the wave number $\underset{\sim}{k}$ by

$$\omega = \underset{\sim}{b}\cdot\underset{\sim}{k} / \sqrt{4\pi\mu\rho_o} \quad . \tag{14}$$

The group velocity for this wave is then

$$\frac{\partial\omega}{\partial\underset{\sim}{k}} = \underset{\sim}{B} / \sqrt{4\pi\mu\rho_o} \quad . \tag{15}$$

We thus see that in this wave, the group velocity is parallel to the direction of the constant magnetic field. Thus, the energy is propagated along the field lines and there is no geometrical spreading of the energy. In addition, the thermodynamic state of the medium is unchanged by the passage of the wave since the density and, therefore, the pressure remain constant.

The relation between the disturbance in velocity v_3^* and the disturbance in the magnetic field is

$$v_3^* = - \frac{b_3^*}{\sqrt{4\pi\mu\rho_o}} \quad . \tag{16}$$

Substituting into the energy equation, we obtain

$$\frac{1}{2} \rho_o |v_3^*|^2 = \frac{|b_3^*|^2}{8\pi\mu} \quad . \tag{17}$$

In this wave, the kinetic energy is balanced by the magnetic energy. Energy is thus equally partitioned between the kinetic and magnetic parts, and, since there is no change in the thermodynamic state of the material, there is no variation in the internal energy.

The compatibility conditions for the mode involving the magnetic disturbance and the velocity disturbance in the plane of the propagation vector in constant magnetic field determine the two phase velocities as

$$c_f = \frac{1}{2} \left[C^2 + \frac{B^2}{4\pi\mu\rho_o} + \frac{B_1 C}{(\pi\rho_o)^{1/2}} \right]^{1/2} + \left[C^2 + \frac{B^2}{4\pi\mu\rho_o} - \frac{B_1 C}{(\pi\rho_o)^{1/2}} \right]^{1/2} \tag{18}$$

$$c_s = \frac{1}{2} \left[C^2 + \frac{B^2}{4\pi\mu\rho_o} + \frac{B_1 C}{(\pi\rho_o)^{1/2}} \right]^{1/2} - \left[C^2 + \frac{B^2}{4\pi\mu\rho_o} + \frac{B_1 C}{(\pi\rho_o)^{1/2}} \right]^{1/2} \quad .$$

c_f is, in general, greater than c_s so that a wave traveling with the velocity c_f is termed the fast wave and c_s the slow wave.

Equations (18) for the phase velocity depend on the special geometry. In general, we can write

$$c_f = C_A \left\{ \frac{1}{2} (1 + \beta^2) + \frac{1}{2} [(1 + \beta)^2 - 4\beta \cos^2 \psi]^{1/2} \right\}^{1/2} \tag{19}$$

$$c_s = C_A \left\{ \frac{1}{2} (1 + \beta^2) - \frac{1}{2} [(1 + \beta)^2 - 4\beta \cos^2 \psi]^{1/2} \right\}^{1/2} ,$$

where $C_A = |\underset{\sim}{B}|/(4\pi\mu\rho_o)^{1/2}$, $\beta = C/C_A$, and ψ is the angle between $\underset{\sim}{B}$ and $\underset{\sim}{n}$. From these relations, it is readily established that c_i given by Equation (13) satisfies the inequalities

$$c_s \leq c_i \leq c_f \quad , \tag{20}$$

$$c_s \leq C \quad , \quad c_i \geq C \quad .$$

By direct calculation, it can be shown that the energy equation for the three waves is

$$\frac{1}{2} \rho \, |\underset{\sim}{v}^*|^2 = \frac{1}{2} \frac{C}{\rho} \rho^2 + \frac{|\underset{\sim}{b}^*|^2}{8\pi\mu} \quad . \tag{21}$$

Thus, in all waves, slow, Alfvén or intermediate, and fast, the energy density associated with each of the modes is divided into equal parts between the kinetic energy on one hand and the

magnetic and internal energies on the other. This point is of considerable importance in identifying waves in space.

The above considerations clearly indicate the formal properties of the three waves. An alternative instructive approach to the equations follows from the introduction of new variables. We consider the case in which the initial magnetic field is along the x_3 direction. We then consider a disturbance in which $\rho = 0$, $v_3 = 0$, and $b_3 = 0$. The continuity equation then is

$$\frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} = 0 \quad , \tag{22}$$

and there is no variation of entropy. We then introduce the variables ξ and j

$$\xi = \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \qquad j = \frac{\partial b_2}{\partial x_1} - \frac{\partial b_1}{\partial x_2} \tag{23}$$

$$\xi^* = i(k_1 v_2^* - k_2 v_1^*) \qquad j^* = i(k_1 b_2^* - k_2 b_1^*) \quad ,$$

where ξ is the component of vorticity or spin of the fluid about the line of force and j is the current density associated with the disturbance.

The equations for the propagation of disturbance then reduce to

$$\frac{\partial \xi}{\partial t} = C_A^2 \frac{\partial j}{\partial x_3} \qquad \frac{\partial j}{\partial t} = \frac{\partial \xi}{\partial x_3} \tag{24}$$

$$-\omega \xi^* = C_A^2 k_3 j^* \qquad -\omega j^* = k_3 \xi^* \quad .$$

$$\omega = C_A k_3 \quad , \tag{25}$$

where C_A is the Alfvén velocity. The vorticity about the line of force is propagated along the line of force at the Alfvén wave velocity as a one-dimensional wave. The thermodynamic state of the system remains constant since the variation of entropy and pressure vanish. Associated with the propagation of the vorticity, there is a variation of current density along the line of force. This interpretation of the intermediate wave in terms of a one-dimensional propagation of vorticity is important in considering the propagation of hydromagnetic waves in nature. Because the wave is propagated one-dimensionally, there is no geometrical spreading and, provided that dissipative effects can be neglected, an initial disturbance remains constant in amplitude along a particular line of force.

The other two modes are not separable. The dispersion relation is

$$\varkappa^2 = \frac{(\omega^2 - C^2 k_3^2)(\omega^2 - C_A k_3^2)}{(C_A^2 + C^2)\omega^2 - C_A^2 C^2 k_3^2} \quad , \tag{26}$$

where $\varkappa^2 = k_1^2 + k_2^2$. The propagation is thus symmetrical about the x_3 axis.

For specified values of the wave number k at frequency ω, Equation (26) determines $\varkappa^2$. The value thus determined might be negative; but $\varkappa$ must be real since the system is conservative and assumed stable. The values of frequency and wave number resulting in a negative $\varkappa^2$ do not give an allowed eigen solution. The region in the ω - k^3 plane of permitted waves is shown in Figure 1. The upper lefthand region of the diagram is bounded by the line $\omega = C_A\, k_3$. Over this region, $\varkappa^2 > 0$. A second region of allowed solutions is bounded by the lines $\omega = C\, k_3$ and $\omega = k_3\, C_A\, C / \sqrt{C_A^2 + C^2}$. Along this line, the wave number $\varkappa$ becomes infinite. This is interpreted physically as the vanishing wave length.

The diagram in Figure 1 is scaled so that $\beta = C/C_A = 0.5$. The upper lefthand region corresponds to the propagation in which the two-dimensional divergence of the plane perpendicular to $\underset{\sim}{B}$ is propagated at a speed c_f. The narrow region corresponds to the propagation of the slow wave. The lines bounding the regions of allowed solution are straight, since the dispersion relation is homogeneous in frequency and wave number; both modes are propagated in a uniform medium without dispersion. It should be noted that the line $\omega = C_A\, k_3$ corresponds to the intermediate or Alfvén wave which has been separated out of the dispersion relation by the choice of variables.

The slowness surface corresponding to the diagnostic diagram of Figure 1 is shown for a fixed frequency ω in Figure 2. The

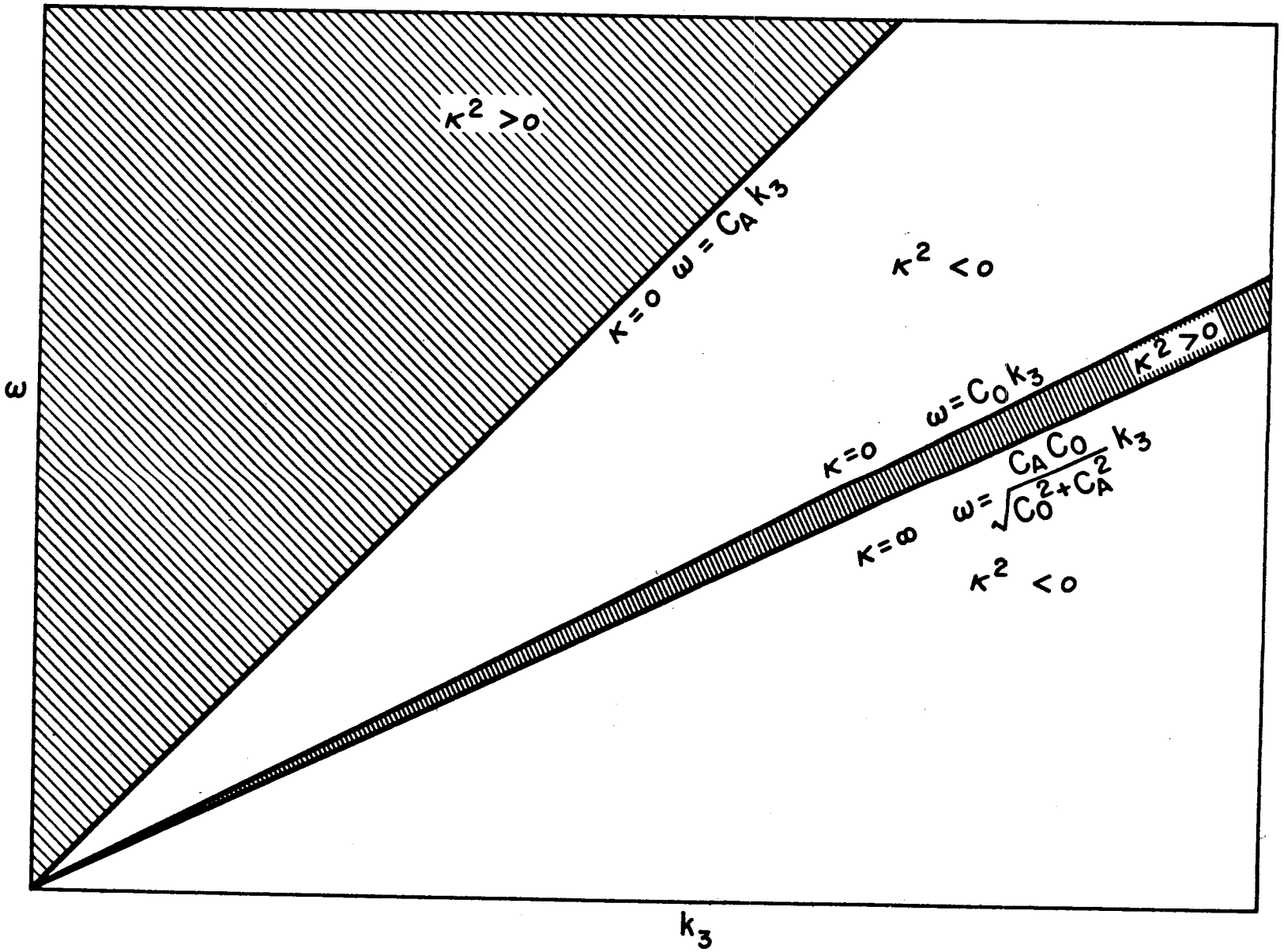


Figure 1: Diagnostic diagram for hydromagnetic wave propagation in a uniform fluid. The diagram is scaled so that $C/C_A = 0.5$. In the upper shaded regions waves are propagated with phase velocity c_f. In the lower narrow band waves are propagated with phase velocity c_s.

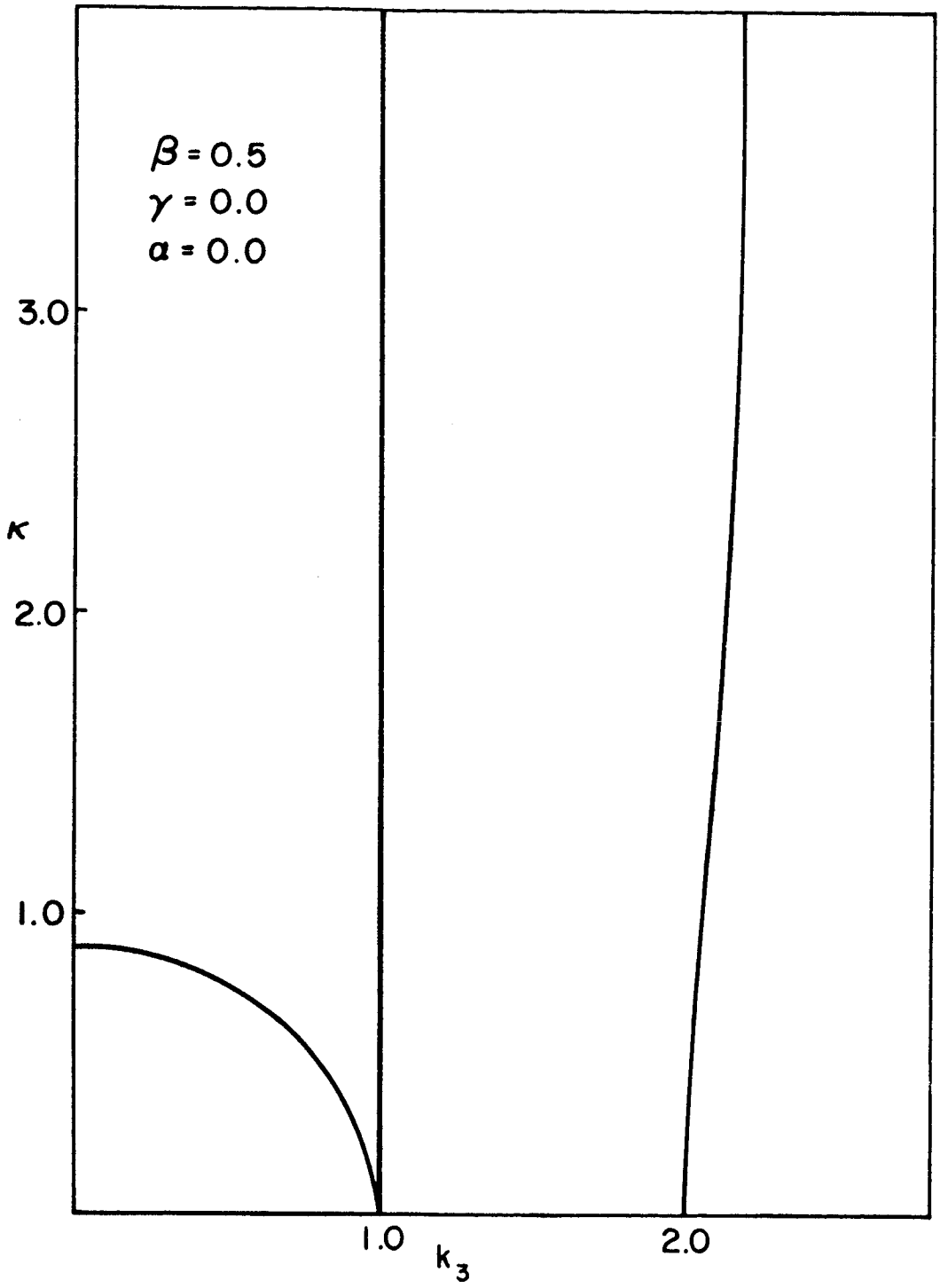


Figure 2: Slowness surfaces for propagation of hydromagnetic waves in a uniform medium. The complete surfaces are obtained by rotation about the k_3 axis and reflection across the $k_3 = 0$ plane. The unit of wave number is ω/C_A. The lines of force are parallel to the k_3 direction. The near-sphere corresponds to wave propagation at phase velocity c_f. The plane corresponds to waves propagated at phase velocity c_i, while the waves along the plane with a bump are propagated at a velocity c_s.

complete surfaces are obtained by rotating the k_3 - $\varkappa$ plane about the k_3 axis and reflecting across the plane $k_3 = 0$. The two sheets of the slowness surface determined by the dispersion relation of Equation (26) are a sphere flattened along the $\varkappa$ direction and a near plane with a bump toward the origin. In addition, the plane representing the one-dimensional propagation of the vorticity has been added. The near sphere corresponds to the nearly spherical spreading mode propagating at a phase velocity c_f.

The surfaces of constant phase can be constructed from the slowness surface. Figure 3 shows the wave form with unit time for a source located at the origin and time $t = 0$, and for $\beta = 0.6, 0.8$. The magnetic lines of force are taken along the x_3 axis and there is cylindrical symmetry about the x_3 axis and across the x_2 - x_3 plane. The unit of distance is $C_A t$. The slowness surface for the intermediate mode degenerates to a point on the x_3 axis since this mode is propagated one-dimensionally. The fast mode represented by the sphere flattened along the $\varkappa$ direction in wave number space maps into a sphere flattened in the x_3 direction. The inverse property of the slowness surface is illustrated in the wave surface for the fast mode. The phase velocity in the direction perpendicular to the lines of force is greater than the velocity along the lines of force.

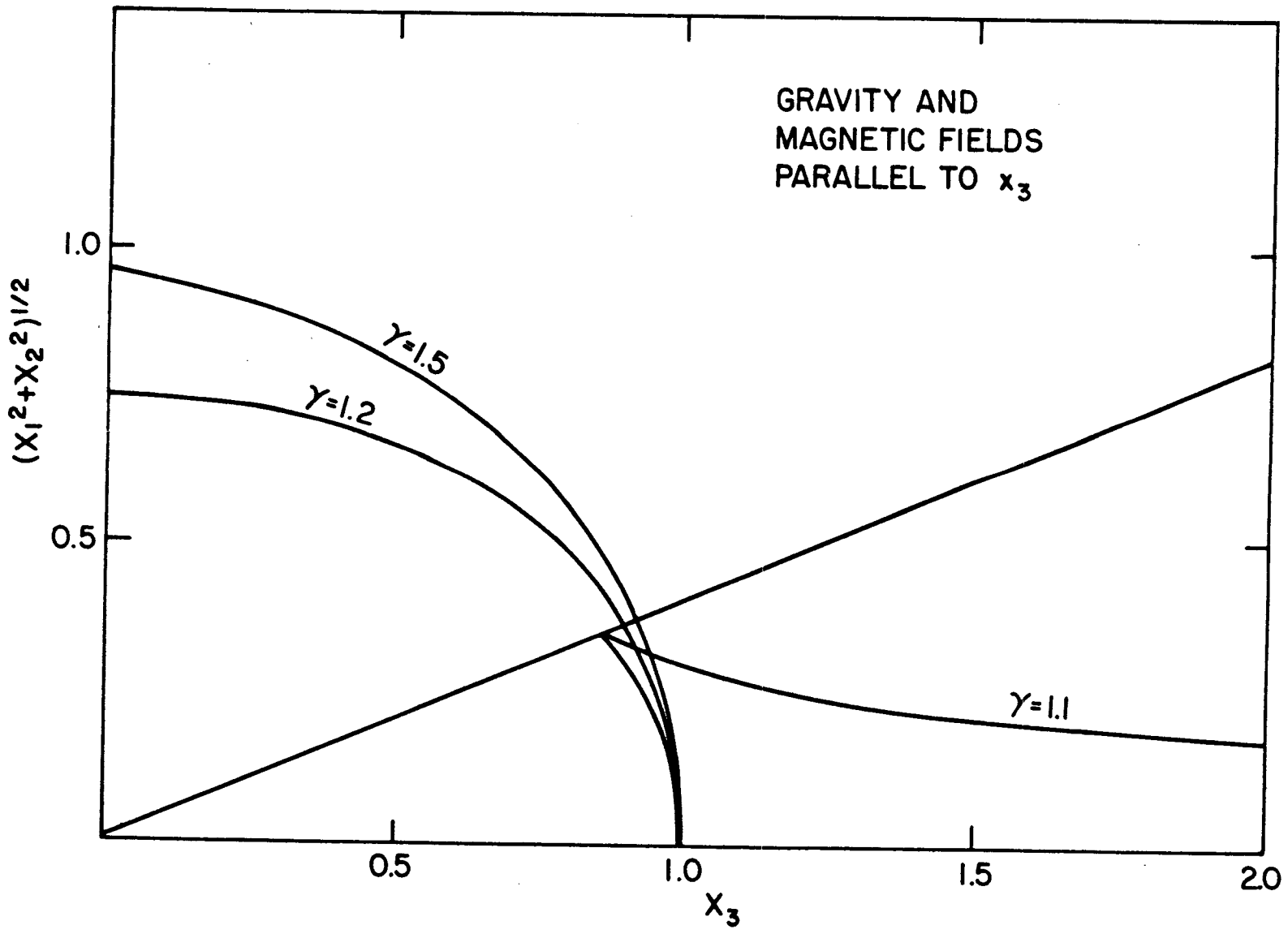


Figure 3: Surfaces of constant phase for a uniform medium. The complete surfaces are obtained by rotation about the x_3 axis and reflection across the x_1 - x_2 plane. The cusped triangles are propagated at velocity c_s; the flattened sphere at a velocity c_f.

The slow mode is propagated almost one-dimensionally along the line of force. The surface of constant phase has the form of a cusp with a conical point on the x_3 axis. This cusp surface is subtended by a cone at the vertex of the origin. The angular bounds of the cone increase with increasing β up to $\beta = 1$.

The discussion above has been based on the assumption that the undisturbed medium is everywhere continuous. We now consider, in outline, the problem of propagation where a surface of discontinuity separates two different states of the undisturbed medium. The discussion is not completely appropriate since the ionosphere-neutral atmosphere and the magnetosphere-interplanetary space boundaries are gradational rather than sharp. Suppose that one of the three disturbance waves is incident on the surface of discontinuity. The basic problems of the theory are then the determination of the laws of reflection and refraction, the motion of the scattered phase fronts, as well as the strength of the various waves in terms of the strength of the incident wave. The problem has been discussed by Ferraro (1954), Roberts (1955), Bazer (1961), and in considerable detail by Bazer and Hurley (1963).

We first consider the question as to what sort of reflected and refracted waves will be generated if a plane wave of a particular type strikes another phase. A simple, geometrical discussion of the problem has been given by Pridmore-Brown

(1963). Let the magnetic field lie in the plane of incidence; then the intermediate or Alfvén wave involves the vorticity about the line of force and, thus, motion perpendicular to the plane of incidence. The other two waves involve motion in the plane.

Suppose the external magnetic field is parallel to the interface, but discontinuous across it. In order that the total pressure be continuous, the fluid pressure must also be discontinuous across the interface. A fast or slow wave is incident on the discontinuity. The angle between the propagation vector and the perpendicular to the plane of incidence is φ. Such a wave will set up a disturbance or ripple in the interface which will propagate at the speed equal to the trace velocity of the wave. The trace velocity is $c_f/\sin\varphi$ or $c_s/\sin\varphi$ depending on whether the incident wave is the fast or slow wave respectively. A disturbance in the plane of the discontinuity will then generate reflected and refracted waves at angles such that their trace velocities match the ripple speed. It is clear that the incident wave will give rise to a reflected wave of the same species at an angle of reflection equal to the angle of incidence. We further suppose that on both sides of the discontinuity the Alfvén velocity is large compared with the sound wave velocity. If the trace velocity on the discontinuity due to the incident wave lies above the Alfvén wave velocity for the second medium, then only the fast wave is transmitted. This can be seen by reference to the diagnostic diagram shown in

Figure 1. If the trace velocity lies between the sound velocity C and $C_A/(C^2 + C_A^2)^{1/2}$, then the slow wave will be propagated into the second medium. If the trace velocity lies outside the shaded regions given in the diagnostic diagram for the second medium, then no wave will be transmitted and the incident wave will be totally reflected. Finally, we note that in the case discussed with the magnetic field parallel to the discontinuity, an intermediate or Alfvén wave will not give rise to a transmitted or reflected wave, since the energy is propagated parallel to the lines of force and no disturbance is set up in the surface of the discontinuity.

We next consider the case where the magnetic field is perpendicular to the interface. In this case, the magnetic field and the pressure must be continuous across the discontinuity. In general, an incident slow or fast wave will generate two reflected and two transmitted waves; in addition to the continuity of pressure and of the normal component of fluid velocity, one must also require continuity of the tangential component of fluid velocity. This last requirement is necessary to ensure continuity of the tangential electric field. An incident Alfvén or intermediate wave gives rise simply to a single reflected and transmitted wave of the same kind.

If the restriction that the magnetic field lies in the plane of incidence is dropped, then all three types of waves can be excited by a single incident wave of any species at an

interface. Mathematically, this is because there are three boundary conditions involving the continuity of pressure and of both tangential components of velocity, and three waves are needed to satisfy these.

Observation of Hydromagnetic Waves at the Earth's Surface

The magnetic field as measured at the earth's surface with sensitive magnetometers having response times on the order of 1/10 of a second shows a bewildering variety of variations. Included among these variations are almost periodic variations of small amplitude. These are usually called micropulsations. Micropulsations vary between day and night and show great variability in duration. Certain micropulsations can be correlated from one geographic locality to another but others are completely uncorrelated. Some micropulsations show a high degree of correlation with geophysical phenomena of various sorts while others do not. Comprehensive descriptions of micropulsations exist yet their origin remains very much a problem.

Geomagnetic pulsations have amplitudes that range from a fraction of a gamma to, on rare occasions, as much as a few tens of gamma. The periods range from 0.1 seconds to 10 minutes. Many of the observations have been made over the low-frequency end of this range, while relatively little is known about the pulsations at high frequencies. It is customary to divide micropulsations into three broad categories: P_c, P_t, and P_g. P_c oscillations were discovered in the records of a very

sensitive horizontal coil at Eskdalenuir (Holmberg, 1953). P_c are a series of pulsations lasting for many hours with periods usually in the range of 10 to 60 seconds and amplitudes on the order of 0.1 gamma. The maximum occurrence frequency is during the morning hours. Well-defined oscillations with periods less than 10 seconds have been observed by Troitskaya (1961) and Benioff (1960). Holmberg found long trains of pulsations during the day, but at night short trains of damped pulsations. Similar results were found by Troitskaya in which she found pulses in the evening, and continuous pulsations in the morning. A study of Jacobs and Sinno (1960) shows a geographical variation in the amplitude of the P_c with the highest amplitude in the auroral regions and the amplitudes continuously decreasing to the equator. Various attempts to measure the Q of these oscillations indicate a value of around 5.

The P_t appear as several series of oscillations, each series usually lasting from 10 to 20 minutes, but the whole series of oscillations no more than an hour. They are well damped, and have a longer period than P_c (from 40 seconds to a few minutes). They have amplitudes on the order of one gamma and at a given station can occur about once every three days. They are found over large areas at low and middle latitudes. As with the P_c, the amplitude variation is such that the amplitudes are greater in the auroral regions and decrease toward the equator. It is generally thought that the P_t's occur simultaneously over large fractions of the earth's surface. The

pulsations are frequency associated with the magnetic disturbances known as magnetic bays.

P_g are the giant micropulsations. These have amplitudes on the order of 20 gamma, and are observed only in the auroral zone and are infrequent at any given station. They are observed most frequently around midnight. The period is longer than that of the P_c, up to several minutes, and the duration can be as much as an hour or more. The Q of the giant pulsation is on the order of 10 or less. The extent of these pulsations in longitude and the variation of their phase with position are as yet poorly determined.

A particularly good description of the giant pulsations has been given by Sugiura (1961). He found damped waves of a period of several minutes with amplitudes up to a hundred gamma in the geomagnetic field in the auroral zone. These damped waves occur simultaneously in magnetically conjugate regions in the northern and southern hemisphere. The waves are nearly elliptically polarized in the plane perpendicular to the local magnetic field.

The above description and elementary classification of geomagnetic pulsations is far from being satisfactory since the time records of various kinds of pulsations grade into each other. Attempts have been made to subdivide P_c into groups based on their frequency range. In addition, often the

micropulsations beat, with the result that the record gives the appearance of pearls on a string with the beating period ranging from seconds to tens of seconds. An alternative scheme for discussing micropulsations has been developed by Campbell. Campbell (1962, 1963) suggests that the oscillations be classified into four period ranges: 0.3 - 4 seconds, 5 - 30 seconds, 40 - 240 seconds, and 240 - 600 seconds. The high-frequency oscillations show a daytime maximum occurrence at high and middle latitudes and there is some correspondence between the magnetic records and geomagnetically conjugate locations. This type of oscillation has not been observed in the equatorial region and the periods and amplitudes are larger toward higher latitudes. The period range 5 - 30 seconds includes some P_c. The time of frequent occurrence is between 1000 and 1400 local time in middle latitudes and about noon at the auroral zone. The amplitude increases toward higher latitudes. The 40 - 240 second frequency range corresponds closely to P_g. Frequent occurrences are about noon in middle latitudes and about midnight in the auroral zone; to date, none have occurred in the equatorial zone.

The low-frequency waves correspond to P_g. As noted above, the vector diagrams show elliptical or circular polarization with a clockwise rotation in the northern hemisphere and a counterclockwise rotation in the southern hemisphere.

The theoretical interpretation of the geomagnetic pulsations is still in a most primitive state. Most workers agree that they represent hydromagnetic waves excited in the magnetosphere, propagated through the magnetosphere at hydromagnetic velocities, and converted at the lower ionospheric boundary into electromagnetic fluctuations which then propagate through the neutral atmosphere. Because of the often almost-sinusoidal character of the pulsations, it is generally thought that they may involve a resonance phenomenon. In addition, the simultaneity of certain of the P_t oscillations suggest that they may be standing waves since the phase does not vary with the position.

The most detailed interpretation of the micropulsations has been on the giant micropulsations. The one-dimensional character of the Alfvén or intermediate mode raises the possibility that this mode may be trapped along a line of force and give rise to a standing wave. Since the charged particle density varies continuously, the mode will be a leaking mode with energy passing through and radiating out as an electromagnetic wave at the end of the line of force. The density contrast in the ionosphere then suggests a Q for this mode of oscillation on the order of 10, all internal losses being neglected (MacDonald, 1961).

The direct calculation of the free period of this mode of oscillation is straightforward. The equation for the

propagation of vorticity satisfies a one-dimensional wave equation. By analogy with the oscillations of a weighted spring it is then possible to calculate the free modes for various models of the charged particle distribution and for a dipole field. Results of these computations are shown in Figure 4. The high-density model corresponds approximately to the charged particle distribution during the solar maximum while the low-density model corresponds to the solar minimum. In the auroral regions, the period for this mode of oscillation varies between 200 and 500 seconds, corresponding approximately to the observed periods. The identification of the giant pulsations as trapped free modes of the intermediate wave thus seems reasonable. As yet, the geographical distribution of the period of the giant pulsations is not well enough known to determine whether or not the variation falls along the curve shown in Figure 4.

Hydromagnetic Waves in Space

A number of workers have interpreted fluctuations of the magnetic field in space as evidence of hydromagnetic waves. The interpretation of such fluctuations is not at all straightforward, particularly in the interplanetary medium. In the magnetosphere, the orientation of the undisturbed magnetic field can be determined with some precision and the disturbances measured relative to the zeroth order field. In the interplanetary medium, the field is carried or generated by the fluctuating streaming plasma and it is difficult to define a

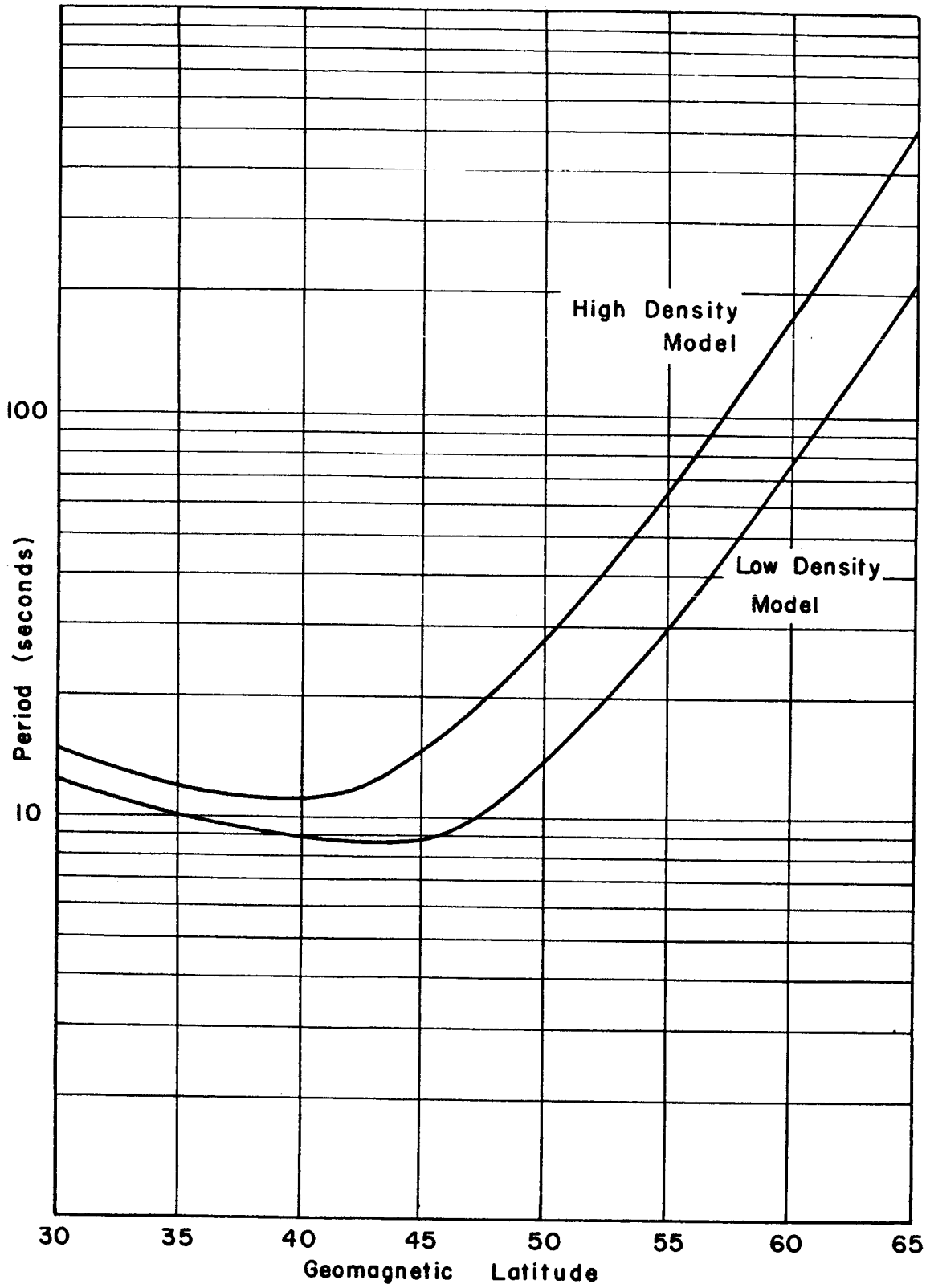


Figure 4: Variation of free period of oscillation with geomagnetic latitude. The high-density model corresponds roughly to conditions during solar maximum, the low-density to solar minimum.

steady field. In order to adequately define a hydromagnetic wave, it is necessary to determine the variation of three components of the magnetic field so that the polarization characteristics can be fixed. The first in situ observations of fluctuation in the distant geomagnetic field interpreted to have been produced by hydromagnetic waves were obtained by Sonett et al. (1960) with a magnetometer aboard the space probe, Pioneer 1. Further observations were reported by Sonett et al. (1962).

Judge and Coleman (1962) observed certain low-frequency hydromagnetic fluctuations within the magnetosphere using magnetometers aboard Explorer 6. The fluctuation exhibited periods between 100 and 500 seconds. Within this range of periods, variations on the order of 10 gamma in amplitude were observed in the measured component of the field. Judge and Coleman interpreted these results as hydromagnetic disturbances propagating in the intermediate mode. The elliptical polarization and the fact that the rotation of the polarization was clockwise to an observer looking along the unperturbed field line toward the north geomagnetic pole showed the similarity of these oscillations to those observed by Sugiura and others in the auroral regions. The transverse oscillations identified by Judge and Coleman could very well be the giant pulsations.

Fluctuations have been reported by Hepner et al. from studies on board Explorer 10 and by Coleman et al. from

magnetometers aboard Mariner 2. In both of these cases, the data reduction has not proceeded far enough to allow a clear interpretation of the results in terms of hydromagnetic waves. These observations are of great interest because of the associated fluctuations in plasma density. Hydromagnetic waves should show changes in the spectrum accompanying the changes in the magnetic field so that the energy balance expressed by Equation (21) is maintained.

While it is fairly certain that the geomagnetic fluctuations observed both at the earth's surface and in space represent hydromagnetic waves, the origin of these waves remains in doubt. It is generally thought (Dungey, 1963) that the disturbances originate at the boundary of the magnetosphere. It is expected that the flow of the interplanetary gas over the surface will generate waves in the same way that wind generates waves on water. As the plasma streams past the bounding surface of the magnetosphere at velocities greater than the local Alfvén velocity, waves can be generated in the surface of the magnetosphere either through shear instability or by fluctuations in the dynamic pressure. The actions of the fluctuations in dynamic pressure can be visualized in a straightforward fashion. We suppose that there is a pressure fluctuation with a certain lengthscale L. This fluctuation will tend to produce longitudinal fast waves of the same scale L traveling at various angles to the solar wind. In general, the wind speed will be different from the phase velocity of the waves. However, if the component of the wind speed

resolved in the direction of the wave just equals the velocity of the wave, then there is a matching between the fluctuation in pressure and the mode of oscillation. A kind of resonance takes place and the wave amplitude builds up. The local Alfvén velocity at the bounding surface is of the order of one-tenth of the streaming velocity of the plasma. Since the rate of build-up is greatest when the resolved component velocity equals the phase velocity, we should expect that the fluctuations in pressure would be most effective on the circular region of the bounding cavity, where the circular region subtends an angle of about 10 degrees.

Waves generated along the bounding surface will be transmitted inwardly with the fast mode propagating spherically while the intermediate mode is guided along the lines of force. Both of these modes are partially trapped within the magnetosphere. These modes then pass into the region where the density of neutral particles becomes so high that collisions of the ionized particles with the neutral particles remove energy from the wave at a high rate (Dessler, 1959; Francis and Karplus, 1960; Karplus et al., 1962). However, part of the energy can be converted into electromagnetic energy and propagated through the earth's neutral atmosphere as ordinary electromagnetic waves.

Observations at ground level can provide some estimate of the intensity of hydromagnetic waves in the upper atmosphere. For example, Ness et al. (1962), in a power spectral analysis

of geomagnetic fluctuations at Fredericksburg, Virginia, found an energy of 0.1 erg cm^{-3} in the frequency interval 0 - 0.5 c/s. The spectrum peaks at about 0.1 c/s with a Q of about 6. Similar values for Q for giant pulsations have been obtained by Sugiura (1961). Since Q is related to the rate of energy dissipation, the quiet time rate at which energy is dissipated in the frequency range of 0 to 0.5 c/s is then of the order of 1×10^{-14} erg cm^{-3} sec^{-1}. With an Alfvén wave velocity of 10^9 cm sec^{-1}, the energy flux into the ionosphere due to hydromagnetic wave dissipation in this frequency range is 10^{-5} erg cm^{-2} sec^{-1}. This energy is completely negligible in considering the total energy budget of the ionosphere.

REFERENCES

Bazer, J., Reflection and refraction of weak hydromagnetic discontinuities, Research Report MH-11, Division of Electromagnetic Research, Courant Institute of Mathematical Sciences, New York University, 1961.

Bazer, J., and O. W. Fleischman, Propagation of weak hydromagnetic discontinuities, Report MH-10, Division of Electromagnetic Research, Courant Institute of Mathematical Science, New York University, 1959a.

Bazer, J., and O. W. Fleischman, Propagation of weak hydromagnetic discontinuities, Physics of Fluids, 2, 366-378, 1959b.

Bazer, J., and J. Hurley, Geometrical hydromagnetics, J. Geophys. Res., 68, 147-174, 1963.

Benioff, H., Observations of geomagnetic fluctuations in the period range 0.3 to 120 seconds, J. Geophys. Res., 65, 1413-1422, 1960.

Bonetti, A., H. S. Bridge, A. J. Lazarus, B. Rossi, and F. Scherb, Explorer X plasma measurements, J. Geophys. Res., 68, 4017-4064, 1963.

Campbell, W. H., Geomagnetic micropulsations, Paper presented at Second Western National Meeting, American Geophysical Union, Stanford, Dec. 27-29, 1962.

Campbell, W. H., Natural electromagnetic field fluctuations in the 3.0 to 0.02 c/s range, Proceedings I.E.E.E., In press, 1963.

Coleman, P. J., L. Davis, E. J. Smith, and C. P. Sonett, Interplanetary magnetic fields, Science, 138, 1099-1100, 1962.

Dessler, A. J., Ionospheric heating by hydromagnetic waves, J. Geophys. Res., 64, 397-401, 1959.

Dungey, J. W., The structure of the exosphere or adventures in velocity space, in Geophysics, Edited by C. DeWitt, J. Hieblot, and A. Lebeau, Gordon and Breach, London, pg. 505-552, 1963.

Fejer, J. A., Hydromagnetic reflection and refraction at a fluid velocity discontinuity, Physics of Fluids, 6, 508-512, 1963.

Ferraro, B. C. A., On the reflection and refraction of Alfvén waves, Astrophysical Journal, 119, 2, 393-406, 1954.

Francis, W. E., and R. Karplus, Hydromagnetic waves in the ionosphere, J. Geophys. Res., 65, 3593-3600, 1960.

Gringauz, K. I., The structure of the earth's ionized gas envelope based on local charged particle concentrations measured in USSR, Space Research, Proc. Intern. Space Sci. Symp., 3rd, Washington, North-Holland Publishing Company, Amsterdam, 574-591, 1962a.

Gringauz, K. I., Some results of experiments in interplanetary space by means of charged particle traps on Soviet space probes, Space Research, Proc. Intern. Space Sci. Symp., 3rd, Washington, North-Holland Publishing Company, Amsterdam, 539-553, 1962b.

Hepner, J. P., N. F. Ness, C. S. Scearce, and T. L. Skillman, Explorer X magnetic field measurements, J. Geophys. Res., 68, 1-46, 1963.

Herlofson, N., Magneto-hydrodynamic waves in a compressible fluid conductor, Nature, 165, 1020-1021, 1950.

Holmberg, E. R. R., Rapid periodic fluctuations of the geomagnetic field, Mon. Not. Roy. Astron. Soc., Geophysical Supplement, 6, 476-481, 1953.

Jacobs, J. A., and K. Sinno, World-wide characteristics of geomagnetic micropulsations, J. Geophysics, 3, 333-353, 1960.

Judge, D., and P. J. Coleman, Observation of low-frequency hydromagnetic waves in the distant geomagnetic field; Explorer VI, J. Geophys. Res., 67, 5071-5090, 1962.

Karplus, R.,W. E. Francis, and A. J. Dragt, Attenuation of hydromagnetic waves in the ionosphere, Planetary and Space Sciences, 9, 771-786, 1962.

Lighthill, M. J., Studies on magneto-hydrodynamic waves and other anisotropic wave motion, Philosophical Transactions, Royal Society, London A, 252, 397-430, 1960.

MacDonald, G. J. F., Spectrum of hydromagnetic waves in the exosphere, J. Geophys. Res., 66, 3639-3670, 1961.

Ness, N. F., T. L. Skillman, C. F. Scearce, J. P. Hepner, Magnetic field fluctuations on the earth and in space, International Conference on Cosmic Rays in the Earth Storm, Part II, Journal of the Physical Society of Japan, 17, Supplement, A2, 27-33, 1962.

Neugebauer, N., and C. W. Snyder, Solar plasma experiments, Science, 138, 1095-1096, 1963.

Nicolet, M., La constitution et la composition de l'atmosphére superiéure, in Geophysics, The Earth's Environment, Edited by C DeWitt, J. Hieblot, and A. Lebeau, Gordon and Breach, London, 201-280, 1963.

Pridmore-Brown, D. C., Reflection and refraction of magneto-acoustic waves of an interphase, Physics of Fluids, 6, 803-805, 1963.

Roberts, P. H., On reflection and refraction of hydromagnetic waves, Astrophysical Journal, 121, 720-730, 1955.

Sonett, C. P., D. L. Judge, A. R. Simms, and J. M. Kelso, A radial rocket survey of the distant geomagnetic field, J. Geophys. Res., 65, 55-68, 1960.

Sonett, C. P., A. R. Simms, and I. J. Abrahms, The distant geomagnetic field, 1. Infinitesimal hydromagnetic waves, J. Geophys. Res., 67, 1191-1208, 1962.

Sugiura, M., Evidence of low-frequency hydromagnetic waves in the exosphere, J. Geophys. Res., 66, 4087-4095, 1961.

Synge, J. L., Elastic waves in anisotropic media, J. Mathematics and Physics, 35, 323-334, 1957.

Troitskaya, V. A., Pulsation of the earth's electromagnetic field with periods of 1 to 15 seconds and their connection with phenomena in the high atmosphere, J. Geophys. Res., 66, 5-18, 1961.

van de Hulst, H. C., Interstellar polarization in magneto-hydrodynamic waves, in Problems of Cosmo Aerodynamics, pp. 46 (International Union Theoretical and Applied Mechanics and International Astronomical Union Proceedings Symposium, Paris, 1949), Published by Central Air Documents Office, Dayton, Ohio, 1951.